

**TRANSMITTAL LETTER TO THE UNITED STATES
DESIGNATED/ELECTED OFFICE (DO/EO/US)
CONCERNING A FILING UNDER 35 U.S.C. 371**

U.S. APPLICATION NO. (if known, sect. 37 C.F.R. 1.5)

09/743533

"EXPRESS MAIL" MAILING CERTIFICATION

Number: EL659498785US Date of Deposit: January 10, 2001

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, BOX PCT, Washington, D.C. 20231, on January 10, 2001.

Typed or Printed Name: Vincent Diaz

Signed: Vincent M. Diaz

INTERNATIONAL APPLICATION NO.

PCT/AU99/00563

INTERNATIONAL FILING DATE

12 July 1999

PRIORITY DATE CLAIMED

10 July 1998

TITLE OF INVENTION

MODIFIED PROTEINS

APPLICANT(S) FOR DO/EO/US

(1) Rudi APPELS, Aranda, Australia; (2) Matthew MORELL, Aranda, Australia;
(3) Frank BEKES, Bectroft, Australia; and (4) Laszlo TAMAS, Budapest, Hungary

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
 - a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☐ has been transmitted by the International Bureau.
 - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☒ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
 - a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☐ have been transmitted by the International Bureau.
 - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
 - d. ☐ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern other document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
 - ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information.

Amendment (under Article 34) and substitute pages 3, 30, 43 - 45
International Preliminary Examination Report
Preliminary Amendment with appended Claims for Examination

17. <input checked="" type="checkbox"/> The following fees are submitted:				CALCULATIONS PTO USE ONLY	
Basic National Fee (37 CFR 1.492(a)(1)-(5)):					
Search Report has been prepared by the EPO or JPO \$860.00					
International preliminary examination fee paid to USPTO (37 CFR 1.482) \$690.00					
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)) \$710.00					
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$1000.00					
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) \$100.00					
ENTER APPROPRIATE BASIC FEE AMOUNT =				\$ 1,000.00	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input checked="" type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				\$ 130.00	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	27	-20 =	7	X \$ 18.00	\$ 126.00
Independent Claims	3	-3 =	0	X \$ 80.00	\$ 0.00
Multiple dependent claims (if applicable)				+ \$ 270.00	
TOTAL OF ABOVE CALCULATIONS =				\$ 1,256.00	
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).					
SUBTOTAL =				\$ 1,256.00	
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				+	
TOTAL NATIONAL FEE =				\$ 1,256.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property					
TOTAL FEES ENCLOSED =				1,256.00	
				Amount to be refunded	\$
				charged	\$
a. <input checked="" type="checkbox"/> A check in the amount of \$1,256.00 to cover the above fees is enclosed.					
b. <input type="checkbox"/> Please charge my Deposit Account No. <u>06-1300</u> in the amount of \$ _____ to cover the above fees.					
c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. <u>06-1300 (Order No. A-70233/RET)</u> .					
NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.					
SEND ALL CORRESPONDENCE TO:					
Richard F. Trecartin, Esq. FLEHR HOHBACH TEST ALBRITTON & HERBERT LLP Suite 3400 Four Embarcadero Center San Francisco, California 94111-4187 Telephone: (415) 781-1989				 SIGNATURE Richard F. Trecartin NAME 31,801 REGISTRATION NUMBER	

09/743533

PATENT

Attorney Docket No.: A-70233/RFT

JC07 Rec'd PCT/PTO 10 JAN 2001

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:APPELS, *et al.*NATIONAL PHASE ENTRY OF:
PCT/AU99/00563

Serial No.: Not Yet Assigned

Filed: Herewith

For: MODIFIED PROTEINS

EXPRESS MAIL CERTIFICATION
UNDER 37 CFR §1.10Express Mail No. EL659498785USDate January 10, 2001

I hereby certify that this document and any documents listed and referred to as enclosed herein are being deposited with the United States Postal Service on the date given above in an envelope bearing "Express Mail Post Office To Addressee" and addressed to: Assistant Commissioner for Patents **Box PCT**, Washington, D.C. 20231.

Typed Name: VINCENT DIAZSigned: Vincent M. Diaz

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents

BOX PCT

Washington, DC 20231

Sir:

Prior to examination, please make the following amendments.

IN THE CLAIMS:

8. (Amended) The method according to [any one of claims 1 to 7] claim 1 wherein the glutenin or seed-storage protein is selected from the group consisting of low molecular weight glutenins, high molecular weight glutenins, gliadins, puroindolines, grain softness proteins, friabilins, and Chloroform/Methanol-soluble proteins.

10. (Amended) A modified glutenin or seed-storage protein having an ability to incorporate into gluten and having an exogenous amino acid domain inserted therein which confers to the modified protein the ability to bind a ligand or other macromolecule produced by the method according to [any one of claims 1 to 9] claim 9.

Serial No.: Not Assigned

Filed: HERewith

15. (Amended) A modified glutenin or seed-storage protein according to [any one of claims 10 to 14] claim 11 wherein the glutenin or seed-storage protein is selected from the group consisting of low molecular weight glutenins, high molecular weight glutenins, gliadins, puroindolines, grain softness proteins, friabilins, and Chloroform/Methanol-soluble proteins.

17. (Amended) An isolated nucleic acid molecule encoding a modified glutenin or seed-storage protein according to [any one of claims 10 to 16] claim 11.

23. (Amended) Use of a modified glutenin or seed-storage protein according to [any one of claims 10 to 16] claim 11 in the preparation of a food product.

25. (Amended) Use of a modified glutenin or seed-storage protein according to [any one of claims 10 to 16] claim 11 in the preparation of a non-food product.

27. (Amended) Use of a grain or part of a grain containing a modified glutenin or seed-storage protein according to [any one of claims 10 to 16] claim 11 in the preparation of a food product.

Serial No.: Not Assigned
Filed: HEREWITH

REMARKS

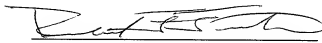
Applicant filed an Article 34 amendment with the International Preliminary Examining Authority, Australia, on May 31, 2000 in response to the Written Opinion. A copy of that amendment is being filed herewith. Applicant has amended claims 8, 10, 15, 17, 23, 25 and 27 of the claims submitted in that amendment to correct claiming multiple dependency. A copy of the claims to be examined in this application are appended hereto.

Respectfully submitted,

FLEHR HOHBACH TEST
ALBRITTON & HERBERT LLP

Date

Jan 10, 2001


Richard F. Trecartin, Reg. No. 31,801

Four Embarcadero Center, Suite 3400
San Francisco, California 94111
Telephone: (415) 781-1989

Serial No.: Not Assigned
Filed: HEREWITH

CLAIMS AS AMENDED FOR EXAMINATION

1. A method of producing a modified glutenin or seed-storage protein, the method comprising adding to the protein an exogenous amino acid domain which confers to the modified protein the ability to bind a ligand or other macromolecule, wherein the modified protein has an ability to incorporate into gluten.
2. The method according to claim 1 wherein the domain includes one or more cysteine residues.
3. The method according to claim 1 wherein one or more cysteine residues are incorporated at one or both ends of the amino acid sequence of the protein.
4. The method according to claim 1 wherein the modified protein further comprises one or more exogenous amino acid residues which confer to the protein an enhanced ability to incorporate into gluten.
5. The method according to claim 1 wherein the ligand or other macromolecule to which the domain binds is selected from the group consisting of lipid and starch.
6. The method according to claim 5 wherein the domain capable of binding lipid is derived from barley oleosin gene or the lipid-binding regions of wheat CM16 protein.
7. The method according to claim 5 wherein the domain capable of binding starch is derived from glucoamylase from *Aspergillus niger*.

Serial No.: Not Assigned
Filed: HEREWITH

8. The method according to claim 1 wherein the glutenin or seed-storage protein is selected from the group consisting of low molecular weight glutenins, high molecular weight glutenins, gliadins, puroindolines, grain softness proteins, friabilins, and Chloroform/Methanol-soluble proteins.

9. The method according to claim 8 wherein the glutenin or seed-storage protein is C hordein from barley.

10. A modified glutenin or seed-storage protein having an ability to incorporate into gluten and having an exogenous amino acid domain inserted therein which confers to the modified protein the ability to bind a ligand or other macromolecule produced by the method according to claim 1.

11. A modified glutenin or seed-storage protein having an ability to incorporate into gluten and having an exogenous amino acid domain inserted therein which confers to the modified protein the ability to bind a ligand or other macromolecule.

12. A modified glutenin or seed-storage protein according to claim 11 wherein the ligand or other macromolecule to which the domain binds is selected from the group consisting of lipid, and starch.

13. A modified glutenin or seed-storage protein according to claim 12 wherein the domain capable of binding lipid is derived from barley oleosin gene or the lipid-binding regions of wheat CM16 protein.

14. A modified glutenin or seed-storage protein according to claim 12 wherein the domain capable of binding starch is derived from glucoamylase from *Aspergillus niger*.

Serial No.: Not Assigned

Filed: HEREWITH

15. A modified glutenin or seed-storage protein according to claim 11 wherein the glutenin or seed-storage protein is selected from the group consisting of low molecular weight glutenins, high molecular weight glutenins, gliadins, puroindolines, grain softness proteins, friabilins, and Chloroform/Methanol-soluble proteins.

16. A modified glutenin or seed-storage protein selected from the group consisting of ANG/SDB/Cys7Cys 236, ANG/OHBD/Cys7Cys236, and ANG/CM16/Cys7Cys126, as hereinbefore defined.

17. An isolated nucleic acid molecule encoding a modified glutenin or seed-storage protein according to claim 11.

18. A cell containing an isolated nucleic acid molecule according to claim 17 such that an expression of the nucleic acid molecule, the cell produces the modified glutenin or seed-storage protein.

19. The cell according to claim 18 selected from the group consisting of bacteria, yeast, plant, insect and mammal.

20. The cell according to claim 19 being *Escherichia coli*.

21. The cell according to claim 19 being *Pichia sp.* or *Saccharomyces cerevisiae*.

22. The cell according to claim 19 being a recombinant wheat cell.

23. Use of a modified glutenin or seed-storage protein according to claim 11 in the preparation of a food product.

Serial No.: Not Assigned
Filed: HEREWITH

24. The use according to claim 23 wherein the food product is selected from the group consisting of leavened or unleavened breads, pasta, noodles, breakfast cereals, snack foods, cakes, pastries, and foods containing flour-based sauces.

25. Use of a modified glutenin or seed-storage protein according to claim 11 in the preparation of a non-food product.

26. The use according to claim 25 wherein the non-food product is selected from the group consisting of films, coatings, adhesives, building materials, and packaging materials.

27. Use of a grain or part of a grain containing a modified glutenin or seed-storage protein according to claim 11.

PATENT
Attorney Docket No.: A-70233/RFT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

APPELS et al.

Serial No. 09/743,533

Filed: January 10, 2001

For: **MODIFIED PROTEINS**

Examiner: Not Yet Assigned

Group Art Unit: Not Yet Assigned

CERTIFICATE OF MAILING

I hereby certify that this correspondence, including listed enclosures, is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to: Assistant Commissioner for Patents, Washington, DC 20231 on: October 1, 2001.

By: Wendy Wilson
Wendy Wilson

AMENDMENT RE SEQUENCE LISTING

BOX SEQUENCE LISTING

Assistant Commissioner for Patents
Washington, DC 20231

Sir:

The present Preliminary Amendment is in response to the Notice to Comply mailed April 30, 2001 and is submitted to comply with requirements for patent applications containing nucleotide sequence and/or amino acid sequence disclosures in adherence with rules 37 C.F.R. § 1.821-1.825. The amendment is submitted within the statutory period, making this a timely response. While Applicants believe that no fees are due at this time, the Commissioner is authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. 06-1300 (Our Order No. A-70233/RFT).

Please amend the application as follows and to comply with requirements for patent applications containing nucleotide sequence and/or amino acid sequence disclosures in adherence with rules 37 C.F.R. § 1.821-1.825 as follows:

Serial No.: 09/743,533
Filed: January 10, 2001

IN THE SPECIFICATION:

Please replace the Sequences Listing starting on page 31 with the enclosed Sequence Listing. Delete the Sequence Listing numbered pages 1-11 at the end of the specification.

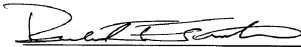
REMARKS

These amendments are made in adherence with 37 C.F.R. § 1.821-1.825. This amendment is accompanied by a floppy disc containing the above named sequence, SEQUENCE ID NUMBERS 1-26, in computer readable form, and a paper copy of the sequence information. The information contained in the computer readable disk is identical to that of the paper copy. This amendment contains no new matter. Applicant submits that this amendment, the accompanying computer readable sequence listing, and the paper copy thereof serve to place this application in a condition of adherence to the rules 37 C.F.R. § 1.821-1.825. Entry of this amendment is respectfully requested. The amendments are made in adherence with 37 C.F.R. § 1.821-1.825.

Respectfully submitted,

FLEHR HOHBACH TEST
ALBRITTON & HERBERT LLP

Dated: October 1, 2001


Richard F. Trecartin, Reg. No. 31,801

Four Embarcadero Center, Suite 3400
San Francisco, California 94111-4187
Telephone: (415) 781-1989

SEQUENCE LISTING

<110> Commonwealth Scientific and Industrial Research
Organisation

<120> Modified Proteins

<130> A-70233/RFT

<140> PCT/AU99/00563

<141> 1999-07-12

<150> AU PP4604

<151> 1998-08-10

<160> 26

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 1

gtcatgaggc aactaaaccc ttgcagccaa gagttgcaat c

41

<210> 2

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 2

ggatccctag accatactcc atatgcatga agcttggtgg gggactgggt g 51

<210> 3

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 3

caagcttgta ccaactccac cgcc 24

<210> 4

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 4

ccatatgcac cgccaggtgt cagtcac 27

<210> 5

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 5
gtcggcaatg aagattgcac c 21

<210> 6
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> unknown

<400> 6
tccaactgcg ttctcctctt ggcc 24

<210> 7
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> unknown

<400> 7
ggatccctag ctccactgag actc 24

<210> 8
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> unknown

<400> 8
 tgcgctcaag ctttaggcaa tgaagattgc acc 33

<210> 9
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> unknown

<400> 9
 cataactccat atgcagctcc actgagactc 30

<210> 10
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> unknown

<400> 10
 caagcttacg atgttgctgg cggg 24

<210> 11
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> unknown

<400> 11
 ccatatgcac cagtaatagc caatagtgc 29

<210> 12
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> unknown

<400> 12
 caaccatgtc ctgaaccttc acc 23

<210> 13
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> unknown

<400> 13
 tggctgttga ggttgcac 18

<210> 14
 <211> 140
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> unknown

<400> 14
 atagaataca gcatgctccc ggccgccatg gccgcgggat tgtcatgagg caactaaacc 60
 cttgcagcgt cccccaacaa gtttcattgca tatggagtat ggtctaggga tccgggtacc 120
 gagctcgaat tcgccctata 140

<210> 15

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> unknown

<400> 15

Met Arg Gln Leu Asn Pro Cys Ser Val Pro Gln Gln Ala Ser Cys Ile

1

5

10

15

Trp Ser Met Val

20

<210> 16

<211> 243

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 16

caagctttga aagccgctac tgcgacagca gctggatcat tgcttgctgt atccgggacta 60
atactagctg gcacagtcac agcactcaca gtggccacac cagtgcctagt catatttagc 120
ccagtgcctag tgccagcggc catagcccta gcgctaattgt cagcaggctt tgtcacgtca 180
ggcgggctgg gcgtggctgc gctgagctcc tttagtgtgt tagccaatac tgctgcata 240
tgg 243

<210> 17

<211> 81

<212> PRT

<213> Artificial Sequence

<220>

<223> unknown

<400> 17

Gln Ala Leu Lys Ala Ala Thr Ala Thr Ala Ala Gly Ser Leu Leu Val
 1 5 10 15
 Leu Ser Gly Leu Ile Leu Ala Gly Thr Val Ile Ala Leu Thr Val Ala
 20 25 30
 Thr Pro Val Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Ala Ile
 35 40 45
 Ala Leu Ala Leu Met Ser Ala Gly Phe Val Thr Ser Gly Gly Leu Gly
 50 55 60
 Val Ala Ala Leu Ser Ser Phe Ser Val Leu Ala Asn Thr Ala Cys Ile
 65 70 75 80
 Trp

<210> 18

<211> 477

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 18

atgaggcaac taaacccttg cagccaagag ttgcaatcac cacaacaatc atatctgccg 60
 cagccatata caaaaaacc atatctaccg caaaaaccat ttccagtgcg gcaaccgttt 120
 cacacacccc aacaatattt cccctatcta ccagaggaat tgtttcccca atatcaaata 180
 ccaaccccc tacaaccaca acaaccatto cccaacaac cacaacaacc tcttccctcg 240
 cccaacaac cattcccttg gcaaccacaa caaccatttc ccagccccca agaaccaatt 300
 cccaacaac cattcccttg gcaaccacaa caaccatttc ccagccccca agaaccaatt 360
 caacaaataa ttttccagca accccaacaa tcataccctg tgcaacctca acagccattt 420
 cctcaacaac ctcaaccagt cccaacaa gcttcatgca tatggagtat ggtctag 477

<210> 19

<211> 158

<212> PRT

<213> Artificial Sequence

<220>

<223> unknown

<400> 19

Met	Arg	Gln	Leu	Asn	Pro	Cys	Ser	Gln	Glu	Leu	Gln	Ser	Pro	Gln	Gln
1				5					10					15	
Ser	Tyr	Leu	Gln	Gln	Pro	Tyr	Pro	Gln	Asn	Pro	Tyr	Leu	Pro	Gln	Lys
		20						25					30		
Pro	Phe	Pro	Val	Gln	Gln	Pro	Phe	His	Thr	Pro	Gln	Gln	Tyr	Phe	Pro
		35					40						45		
Tyr	Leu	Pro	Glu	Glu	Leu	Phe	Pro	Gln	Tyr	Gln	Ile	Pro	Thr	Pro	Leu
	50					55					60				
Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	Pro	Arg
65					70					75				80	
Pro	Gln	Gln	Pro	Phe	Pro	Trp	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro
			85						90					95	
Gln	Glu	Pro	Ile	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro
		100						105						110	
Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Ile	Ile	Phe	Gln	Gln	Pro
		115					120						125		
Gln	Gln	Ser	Tyr	Pro	Val	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro
	130						135						140		
Gln	Pro	Val	Pro	Gln	Gln	Ala	Ser	Cys	Ile	Trp	Ser	Met	Val		
145						150							155		

<210> 20

<211> 338

<212> DNA

<213> *Aspergillus niger*

<400> 20

```

aagcttctac cactcccacc gccgtggtg tgactttcga tctgacagct accaccacct 60
acggcgagaa catctacctg gtcggatcga tctctcagct gggtgactgg gaaaccagcg 120
acggcatagc tctgagtgtc gacaagtaca cttccagcga cccgctctgg tatgtcactg 180
tgactctgcc ggctggtgag tcgttttgagt acaagtttat ccgcattgag agcgatgact 240

```

ccgtggagtg ggagagtgat cccaaccgag aatacacogt tcctcaggcg tgcggaacgt 300
cgaccgcgac ggtgaactgac acctggcggt gcatatgg 338

<210> 21

<211> 112

<212> PRT

<213> *Aspergillus niger*

<400> 21

Ala	Ser	Thr	Thr	Pro	Thr	Ala	Val	Ala	Val	Thr	Phe	Asp	Leu	Thr	Ala
1				5					10					15	
Thr	Thr	Thr	Tyr	Gly	Glu	Asn	Ile	Tyr	Leu	Val	Gly	Ser	Ile	Ser	Gln
			20					25					30		
Leu	Gly	Asp	Trp	Glu	Thr	Ser	Asp	Gly	Ile	Ala	Leu	Ser	Ala	Asp	Lys
		35					40					45			
Tyr	Thr	Ser	Ser	Asp	Pro	Leu	Trp	Tyr	Val	Thr	Val	Thr	Leu	Pro	Ala
		50				55				60					
Gly	Glu	Ser	Phe	Glu	Tyr	Lys	Phe	Ile	Arg	Ile	Glu	Ser	Asp	Asp	Ser
		65			70					75					80
Val	Glu	Trp	Glu	Ser	Asp	Pro	Asn	Arg	Glu	Tyr	Thr	Val	Pro	Gln	Ala
				85					90					95	
Cys	Gly	Thr	Ser	Thr	Ala	Thr	Val	Thr	Asp	Thr	Trp	Arg	Cys	Ile	Trp
				100				105						110	

<210> 22

<211> 371

<212> DNA

<213> Artificial Sequence

<220>

<223> unknown

<400> 22

aagcttttcg caatgaagat tgcaccccat ggtgagtagc tetgatcact ccaactcccaa 60
gtgcccgtga ctatgtggaa caacaagcat gtcgcatcga aacgcccggg tcgccgtacc 120
tcgccaagca gcagtgtctg ggggagcttg caaacattcc gcagcagtcg cgtatgccagg 180

cgctgcgcta cttcatgggg ccgaagtctc gtccggatca gagcggcctc atggaactcc 240
 ccggatgccc tagggagggtg cagatggact tcgtgaggat actcgtcacg ccgggggtact 300
 gcaacttgac gaccgttcac aacctccgt actgcctcgc tatggaggag tctcagtgga 360
 gctgcatatg g 371

<210> 23

<211> 123

<212> PRT

<213> Artificial Sequence

<220>

<223> unknown

<400> 23

Ala	Ile	Gly	Asn	Glu	Asp	Cys	Thr	Pro	Trp	Met	Ser	Thr	Leu	Ile	Thr
1				5					10					15	
Pro	Leu	Pro	Ser	Cys	Arg	Asp	Tyr	Val	Glu	Gln	Gln	Ala	Cys	Arg	Ile
			20					25					30		
Glu	Thr	Pro	Gly	Ser	Pro	Tyr	Leu	Ala	Lys	Gln	Gln	Cys	Cys	Gly	Glu
			35				40					45			
Leu	Ala	Asn	Ile	Pro	Gln	Gln	Cys	Arg	Cys	Gln	Ala	Leu	Arg	Tyr	Phe
			50				55					60			
Met	Gly	Pro	Lys	Ser	Arg	Pro	Asp	Gln	Ser	Gly	Leu	Met	Glu	Leu	Pro
			65			70				75				80	
Gly	Cys	Pro	Arg	Glu	Val	Gln	Met	Asp	Phe	Val	Arg	Ile	Leu	Val	Thr
			85					90					95		
Pro	Gly	Tyr	Cys	Asn	Leu	Thr	Thr	Val	His	Asn	Thr	Pro	Tyr	Cys	Leu
			100					105					110		
Ala	Met	Glu	Glu	Ser	Gln	Trp	Ser	Cys	Ile	Trp					
			115					120							

<210> 24

<211> 123

<212> PRT

<213> Artificial Sequence

<220>

<223> unknown

<400> 24

Ala	Ile	Gly	Asn	Glu	Asp	Cys	Thr	Pro	Trp	Thr	Ser	Thr	Leu	Ile	Thr
1				5					10					15	
Pro	Leu	Pro	Ser	Cys	Arg	Asn	Tyr	Val	Glu	Glu	Gln	Ala	Cys	Arg	Ile
			20					25					30		
Glu	Met	Pro	Gly	Pro	Pro	Tyr	Leu	Ala	Lys	Gln	Glu	Cys	Cys	Glu	Gln
			35				40					45			
Leu	Ala	Asn	Ile	Pro	Gln	Gln	Cys	Arg	Cys	Gln	Ala	Leu	Arg	Tyr	Phe
			50				55					60			
Met	Gly	Pro	Lys	Ser	Arg	Pro	Asp	Gln	Ser	Gly	Leu	Met	Glu	Leu	Pro
65					70					75				80	
Gly	Cys	Pro	Arg	Glu	Val	Gln	Met	Asn	Phe	Val	Pro	Ile	Leu	Val	Thr
				85					90				95		
Pro	Gly	Tyr	Cys	Asn	Leu	Thr	Thr	Val	His	Asn	Thr	Pro	Tyr	Cys	Leu
			100					105					110		
Gly	Met	Glu	Glu	Ser	Gln	Trp	Ser	Cys	Ile	Trp					
			115					120							

<210> 25

<211> 377

<212> DNA

<213> wheat

<400> 25

```

aagcttacga tgttgctggc gggggtggtg ctcaacaatg cctgtagag acaaagctaa 60
attcatgcag gaattacctg ctatgcgat gctcaacgat gaaggatttc ccggtcacct 120
ggcgttggtg gaaatggtgg aagggaggtt gtcaagagct ccttggggag tgttgcagtc 180
ggctcggcca aatgccaccg caatgccgct gcaacatcat ccaggggtca atccaaggcg 240
atctcgggtg catcttcgga ttacagcgtg atcgggcaag caaagtgata caagaagcca 300
agaacctgcc gccacaggtg aaccagggcc ctccctgcaa catccccggc actattggct 360
attactggtg catatggg

```

<210> 26

<211> 125

<212> PRT

<213> wheat

<400> 26

Ala	Tyr	Asp	Val	Ala	Gly	Gly	Gly	Gly	Ala	Gln	Gln	Cys	Pro	Val	Glu
1				5					10					15	
Thr	Lys	Leu	Asn	Ser	Cys	Arg	Asn	Tyr	Leu	Leu	Asp	Arg	Cys	Ser	Thr
			20					25					30		
Met	Lys	Asp	Phe	Pro	Val	Thr	Trp	Arg	Trp	Trp	Lys	Trp	Trp	Lys	Gly
			35				40					45			
Gly	Cys	Gln	Glu	Leu	Leu	Gly	Glu	Cys	Cys	Ser	Arg	Leu	Gly	Gln	Met
			50			55					60				
Pro	Pro	Gln	Cys	Arg	Cys	Asn	Ile	Ile	Gln	Gly	Ser	Ile	Gln	Gly	Asp
65				70					75				80		
Leu	Gly	Gly	Ile	Phe	Gly	Phe	Gln	Arg	Asp	Arg	Ala	Ser	Lys	Val	Ile
			85					90					95		
Gln	Glu	Ala	Lys	Asn	Leu	Pro	Pro	Arg	Cys	Asn	Gln	Gly	Pro	Pro	Cys
			100					105				110			
Asn	Ile	Pro	Gly	Thr	Ile	Gly	Tyr	Tyr	Trp	Cys	Ile	Trp			
			115				120				125				

Modified ProteinsTechnical Field

The present invention is directed generally to producing modified proteins and particularly to producing modified glutenin or seed-storage proteins.

Background Art

Wheat storage proteins are classified on the basis of their solubility into two classes. The gliadins are readily soluble in aqueous alcohols and are monomeric proteins with only intramolecular disulphide bonds. The glutenins are present in high molecular weight polymers, stabilised by intermolecular disulfide bonds and are not soluble in aqueous alcohols without reducing agent (Kasarda 1989). These proteins are present in high amount in the endosperm and are considered to act as a store of nitrogen, carbon and sulphur for seed germination.

Glutenins form a continuous proteinaceous network called gluten. The unique physico-chemical properties of gluten determine the ability of wheat dough to be processed into baked goods (bread, biscuits, cakes), pasta noodles and other food products. It is understood that the glutenins, which form crosslinks with each other through disulfide bonds, are the most important molecules producing the viscoelastic properties of wheat flour dough (MacRitchie 1992). The unique position of wheat in bread making is due to the ability of the dough to retain gas on expansion. The gluten accounts for about 10% of the dough, and consists mainly of proteins (70-80%) together with starch and lipids. Starch could be granular and damaged starch. The lipid reserves of wheat are non-polar, structural and endosperm lipids (Gan et al., 1995). Structural lipids are also called polar lipids. The endosperm lipids are divided into non-starch lipids and starch lysophospholipids. The structure and properties of gluten are determined by molecular interactions and it is important that these be understood if the functional properties of gluten are to be manipulated.

A dough results from a large variety of interactions between flour constituents facilitated by water. Starch takes up about 46% of the water and damaged starch contributes significantly to the water absorption. It has been shown that during hydration, proteins exude visible strands or fibrils. Specific proteins of flour are bound to flour lipids (polar) upon addition of water (Morrison 1989).

Dough development is visualised as a re-orientation of glutenin polymers to form a membrane network with viscoelasticity and gas retaining properties. Covalent (disulfide) and noncovalent (hydrogen, hydrophobic and ionic) bonds are involved in formation of a fully developed dough.

5 Interactions are further modified during fermentation, baking and even after baking. The disulphide bonds of flour proteins play a key role in the interactions in dough. The bonds form relatively strong crosslinks within and between polypeptide chains and also stabilise other less energetic bonds. Disulphide bonds provide the required stability for the protein matrix until

10 the loaf structure is set by the gelatinisation of the starch and the thermal denaturation of the proteins during baking. Hydrogen bonds are considerably weaker than covalent bonds, but contribute significantly to the structure of dough. A unique feature is the ability to interchange with other hydrogen bonds, which facilitate reorientation of protein chains and allow for stress relaxation. Hydrophobic bonds result from nonpolar groups of flour

15 constituents. Because these bonds are reversible, they can readily accommodate viscous flow and thereby facilitate mechanical dough development. Ionic bonds play relatively small part in dough structure formation but some specific components have an ionisable part or parts.

20 Therefore ionic bond interactions could be important for the rheological properties (for a review, see Bushuk 1998).

A major limitation to evaluating the contributions of various groups of proteins, and of specific structural features of these molecules, to dough functionality has been the lack of appropriate systems that allow specific

25 proteins to be incorporated and tested within the dough. The situation has recently changed, however, due to two advances. The first is the development of small scale testing equipment (Mixograph, Extensograph) with appropriate procedures for incorporating exogenous proteins, including polymeric glutenins into the dough (Bekes et al., 1994). Advantages of this

30 system are the small amount of proteins required for test and the ability to rapidly test multiple samples produced by, for example, protein engineering. The second recent advance is the development of a reliable transformation system for wheat (Weeks et al., 1993, Witrzens et al., 1998), which allows the modification of storage protein composition by the expression of new

35 proteins with, for instance, designed characteristics.

ART 34 AMDT

3

To alter protein-protein, protein-lipid and protein-starch interactions within the gluten matrix, the present inventors have developed a system which enables the incorporation of new surface active molecules or parts of molecules into the gluten matrix.

Disclosure of Invention

In a first aspect, the present invention consists in a method for producing a modified glutenin or seed-storage protein, the method comprising adding to the protein an exogenous amino acid domain which confers to the modified protein the ability to bind a ligand or other macromolecule, wherein the modified protein has an ability to incorporate into gluten.

In one preferred embodiment, the modified glutenin or seed-storage protein further contains one or more exogenous amino acid residues added to its amino acid sequence. More preferably, the one or more amino acid residues are one or more cysteine residues. Preferably, the one or more cysteine residues are incorporated at one or both ends of the amino acid sequence of the protein. The addition of the one or more cysteines allows the modified proteins to be more easily incorporated into gluten in use. The further modifications to the glutenin or seed-storage proteins produced according to the present invention allow or can assist in the incorporation of that protein into the gluten network for food or industrial use.

The present inventors have found that incorporating exogenous amino acid sequences (domains) from proteins other than glutenins or seed storage proteins into glutenin or seed-storage proteins modifies the general properties of gluten, particularly when the modified proteins are used in a range of food applications.

Figure 16 provides a schematic of the scheme for identifying transgenes in transformed wheat plants by polymerase chain reaction (PCR). A primer pair straddling the interface between the gene and its promoter (from the gene for the high molecular weight glutenin Bx17 for example) ensures that no false positives are detected arising from the high homology between C hordein and gliadin genes.

The binding domain can be any domain that will bind ligands that may be useful in food preparation or in food compositions. In a preferred form, the binding domain is a ligand capable of binding lipids or starches. The

present inventors have found that the lipid-binding domain of the barley
oleosin gene, the lipid-binding regions of the wheat CM16 protein, and the
starch-binding domain of the glucoamylase from *Aspergillus niger* are
particularly suitable for the present invention. It will be appreciated,
5 however, that other natural or modified domains would also be suitable for
the present invention.

One glutenin or seed-storage protein that has been modified by the
present inventors is the C hordein gene from barley. It will be appreciated,
however, that other glutenin or seed-storage proteins may also be modified
10 according to the present invention. In wheat, such glutenin or seed proteins
include low molecular weight glutenins, high molecular weight glutenins,
gliadins, puroindolines or grain softness proteins (also known as friabilins),
or Chloroform/Methanol-soluble proteins. Homologues of these proteins
exist in other cereals such as diploid, tetraploid and hexaploid wheats, rye,
15 triticale, barley, oats, rice, sorghum, millet and maize and the genes encoding
these proteins may also be modified according to the present invention.

In a second aspect, the present invention consists in a modified
glutenin or seed-storage protein having a domain inserted therein which
confers to the protein the ability to incorporate into gluten or bind a ligand or
20 other macromolecule.

In one preferred embodiment, the modified glutenin or seed-storage
protein is produced by the method according to the first aspect of the present
invention.

In another preferred embodiment, the modified glutenin or seed-
25 storage protein is ANG/SBD/Cys7Cys236, ANG/OHBD/Cys7Cys236 or
ANG/CM16/Cys7Cys236.

In a third aspect, the present invention consists in an isolated nucleic
acid molecule encoding a modified glutenin or seed-storage protein according
to the second aspect of the present invention.

30 In a fourth aspect, the present invention consists in an isolated nucleic
acid molecule according to the third aspect of the present invention
incorporated into a cell such that on expression of the nucleic acid molecule,
the cell produces the modified glutenin or seed-storage protein.

The cell may be a recombinant bacterial cell, for example, which is
35 capable of producing the modified glutenin or seed-storage protein.
Preferably the bacterial cell is *Escherichia coli*. Alternatively, the cell may be

a yeast such as *Pichia* sp. or *Saccharomyces cerevisiae*, an insect cell using an expression system such as the baculovirus expression system, or a mammalian cell. Alternatively, the cell may be a plant cell of a recombinant plant which is capable of producing the modified glutenin or seed-storage protein in the plant's seeds. Preferably the plant cell is a recombinant wheat cell.

In fifth aspect, the present invention consists in the use of a modified glutenin or seed-storage protein according to the second aspect of the present invention in the preparation of a food product.

Examples of food products include leavened or unleavened breads, pasta, noodles, breakfast cereals, snack foods, cakes, pastries or other foods containing flour-based sauces or ingredients.

The modified glutenin or seed-storage proteins according to the present invention, in use, are capable of modifying the structure of doughs and other materials containing gluten in ways which add value and utility to the resultant product. The modified glutenin or seed-storage proteins are suitable for use in the food industry as modifiers of food properties.

The present inventors have shown that modified proteins according to the present invention can be produced in bacterial fermentation and that large scale production of the proteins for commercial use is possible.

In sixth aspect, the present invention consists in the use of a modified glutenin or seed-storage protein according to the second aspect of the present invention in the preparation of a non-food product.

Examples of non-food products include, but not limited to, films, coatings, adhesives, building materials or packaging materials. It will be appreciated that the modified proteins according to the present invention would have the same non-food uses as for normal glutenin or seed-storage proteins.

In a seventh aspect, the present invention consists in the use of a grain or part of a grain containing a modified glutenin or seed-storage protein according to the second aspect of the present invention in the preparation of a food product.

It will be appreciated that the modified glutenin or seed-storage proteins according to the present invention may be contained in, or produced by, a transgenic plant produced by the fourth aspect of the present invention. Thus, the grain or other plant products produced by these plants may be used

in the crude form of flour, semolina, bran, pollard, germ fraction, or the like. Such plants may produce a mixture of normal and modified glutenin or seed-storage proteins which can also be used. Mixtures of modified glutenin or seed-storage proteins and other plant-based materials can also be prepared to form improved gluten.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

In order that the present invention may be more clearly understood, preferred forms will be described with reference to the following examples and drawings.

Brief Description of Drawings

Figure 1: Sequence of the pJANGΔCys7Cys236 vector (SEQ ID NO: 14 and SEQ ID NO: 15). Restriction sites and enzymes of the MCS of pJCKm and pGEM-T plasmids are written in bold. Cloning sites for gene insertion into ANG are underlined and written in bold. Cleavage sites for sub-cloning the gene into pET 11d expression vector are underlined.

Figure 2: Nucleotide (SEQ ID NO: 16) and amino acid sequence (SEQ ID NO: 17) of Oleosin Hydrophobic Binding Domain (OHBD). Arrowheads are indicating the direction of primer extension. Smaller letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Amino acid sequence of the central anti-parallel domain is designated by two arrows.

Figure 3: Nucleotide (SEQ ID NO: 18) and amino acid sequence (SEQ ID NO: 19) of ANGΔCys7Cys236 (molecular mass 18.5 kDa) Cysteine residues are marked with bold letters. Extra two amino acids and six nucleotides are written with smaller letters.

Figure 4: Nucleotide (SEQ ID NO: 20) and amino acid sequence (SEQ ID NO: 21) of the starch binding domain of 1,4-α-D-glucan glucohydrolase from *Aspergillus niger* (molecular mass 11.9 kDa). Small letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Restriction sites are in bold and underlined.

Figure 5: Nucleotide (SEQ ID NO: 22) and amino acid sequences of CM16 (SEQ ID NO: 23) and CM17 (SEQ ID NO: 24) (molecular mass is 13.4

kDa). Small letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Restriction sites are in bold and underlined. First amino acid sequence under the nucleotide sequence represents the CM16 protein, while only differences are shown in CM17 protein.

Figure 6: Nucleotide (SEQ ID NO: 25) and amino acid sequence (SEQ ID NO: 26) of Puroindoline A (molecular mass is 14.3 kDa). Small letters are indicating extra nucleotides for cloning and amino acids from ANG molecule. Restriction sites are in bold and underlined.

Figure 7: Expression of the ANG/OHBD/Cys7Cys236 gene. SDS-PAGE analysis followed by Coomassie blue staining. Lane A contains standard molecular weight markers. Lane B shows a control in which the host *E. coli* cell contains the pET-11d vector. Lane C shows the expression of the recombinant ANG/OHBD/Cys7Cys236 protein (migrating slightly further than the 30 kDa marker) from pET-11d containing the ANG/OHBD/Cys7Cys236 gene.

Figure 8: Expression of the ANG/OHBD/Cys7Cys236 gene. SDS-PAGE analysis followed by Western blot using antibodies to C-hordein. Lane A contained standard molecular mass markers. Lane B contained extract of an *E. coli* cell expressing a modified C-hordein gene containing a single cysteine in the N-terminal region. Lane C contained extract from *E. coli* containing the ANG/OHBD/Cys7Cys236 gene in the pET-11d expression plasmid.

Figure 9: SDS-PAGE analysis of the purified ANG/CM16/Cys7Cys236 gene product by Coomassie blue staining. Lane A contains standard protein molecular weight markers. Lanes B to E show ethanol-soluble extracts of the crude *E. coli* lysates. Lane B contained extract from cells containing the control plasmid, pET-11d. Lanes C, D and E contained ethanol soluble extracts of cells harbouring the pET-11d vector containing the ANG/CM16/Cys7Cys236 gene, prepared from cells 2, 4 and 6 hours, respectively, after induction of protein synthesis using IPTG.

Figure 10: Expression of the ANG/PIN-A/Cys7Cys236 gene. SDS-PAGE gel stained with Coomassie blue. Lane A contains standard protein molecular weight markers. Lane B contained extract of cells harbouring the control pET-11d plasmid, lanes C and D contained extracts of cells harbouring the pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene. Lane C contains extract 2 hours after IPTG induction. lane D contains extract prepared 6 hours following IPTG induction.

Figure 11: Expression of the ANG/PIN-A/Cys7Cys236 gene. SDS-PAGE analysis followed by Western blotting with anti-puroindoline A antibodies. Lanes A and B contained extract of cells harbouring pET-11d containing the ANG/PIN-A/Cys7Cys236 gene, 2 hours and 6 hours after induction respectively. Lane C contains a Western blot of cells harbouring a plasmid containing the puroindoline A gene alone (not inserted into the ANGΔCys7Cys236 gene). Lane D contained extract of cells harbouring the control pET-11d vector.

Figure 12: Expression of the ANG/PIN-A/Cys7Cys236 gene. SDS-PAGE analysis followed by Western blotting with anti-hordein antibodies. Lane A contains molecular weight markers. Lane B contained extract of cells harbouring the control pET-11d vector. Lane C contained extract of cells harbouring a plasmid which contains the puroindoline A gene alone. Lane D contained extract of cells harbouring pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene.

Figure 13: Plasmids for the transformation of wheat with constructs containing novel protein genes. The vectors are based on the plasmid pTLZHMWcas which has been modified by insertion of genes into the KpnI and BamHI site region. The vectors are: pTLZ-ANGCys7Cys13, pTLZ-ANGCys236 and pTLZ_ANG/OHBD/Cys7Cys236.

Figure 14: Expression and β -cyclodextrin affinity purification of ANG/SBD/Cys7Cys236. Native- and SDS-PAGE gels of column fractions from a 5 ml β -cyclodextrin-sepharose affinity column loaded with extracts from IPTG-induced cultures of strain AD494(DE3) bearing either pET11d control plasmid or the pET-SBD plasmid for the expression of ANG/SBD/Cys7Cys236. Panel A = SDS-PAGE of fractions from AD494(DE3)/pET-SBD. Panel B = native PAGE of fractions from AD494(DE3)/pET-SBD. Panel C = SDS-PAGE of fractions from AD494(DE3)/pET11d. Panel D = native PAGE of fractions from AD494(DE3)/pET11d. Lane M = molecular mass markers with masses indicated in kDa. Lane 1 = the unbound flow-through from column loading. Lanes 2-5 = four 5 ml washes with column buffer (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 40 mM dithiothreitol, 1 mM EDTA). Lanes 6-10 = five 5 ml elutions with column buffer containing 15 mM β -cyclodextrin.

Figure 15: Schematic of the method of producing a modified seed storage protein containing a binding domain for a macromolecule. In this example, the C-hordein gene is altered to produce an analogue glutenin (ANG) protein containing cysteine residues in the N- and/or C-terminal domains. This protein is then further modified by replacement of part of its repetitive domain with a binding domain from another protein. The addition of cysteine residues allows for the incorporation of the protein into, for example, the gluten macropolymer in wheat flour. The addition of a specific binding domain provides for the non-covalent linkage of the gluten network to other macromolecules (e.g. starch, lipids). Such a protein may be expressed in recombinant bacteria, yeast, or other heterologous host for use as an ingredient or additive, or it may be produced in the seed of a transgenic plant, for example wheat.

Figure 16: Schematic of the scheme for identifying transgenes in transformed wheat plants by polymerase chain reaction (PCR). A primer pair straddling the interface between the gene and its promoter (from the gene for the high molecular weight glutenin Bx17) ensures that no false positives are detected arising from the high homology between C hordein and gliadin genes.

Figure 17: Agarose gel showing PCR-based identification of a transgenic wheat plant containing a gene for a modified seed storage protein with incorporated ligand-binding domain. In this example, the gene for ANG/OHBD/Cys7Cys236 was detected in the plasmid used for transforming the wheat plant (lane 1) and in a putative transformant wheat plant (lane 2), but is not detected in a non-transformed control wheat plant (lane 3).

Figure 18: Starch-binding activity of ANG/SBD/Cys7Cys236 protein assayed by native-PAGE. Protein from each of fraction 6 and 7 from the β -cyclodextrin affinity column of the urea extract of strain AD494(DE3)/pET-SBD was loaded onto native gels containing starch (panels B and D) and/or β -cyclodextrin (panels C and D) or neither starch nor β -cyclodextrin (panel A). The reduction in electrophoretic mobility in the presence of starch alone (panel B) is indicative of the specific starch-binding activity of each protein.

Figure 19: Demonstration of the oxidative polymerisation of a modified seed-storage protein containing a macromolecule-binding domain. In this example, ANG/SBD/Cys7Cys236 was oxidised to form disulphide-linked polymeric species (indicated by arrows in the scanned image of the

SDS-PAGE gel; lanes 1, 2) in the presence or absence of added β -cyclodextrin. The polymeric species were capable of being subsequently reduced by the disulphide-specific reductant dithiothreitol (DTT; lanes 3, 4), showing that the polymerisation was due to the presence of introduced

cysteine residues at the N- and C-termini of the ANG portion of the molecule. M = molecular weight markers with masses indicated (kDa); Lane 1 = ANG/SBD/Cys7Cys236 oxidised in the presence of β -cyclodextrin; Lane 2 = ANG/SBD/Cys7Cys236 oxidised in the absence of β -cyclodextrin; Lane 3 = ANG/SBD/Cys7Cys236 oxidised in the presence of β -cyclodextrin and then reduced with DTT;

Lane 4 = ANG/SBD/Cys7Cys236 oxidised in the absence of β -cyclodextrin and then reduced with DTT;

Lane 5 = reduced ANG/SBD/Cys7Cys236 in the presence of β -cyclodextrin; Lane 6 = reduced ANG/SBD/Cys7Cys236 in the absence of β -cyclodextrin.

Modes for Carrying Out the Invention

MATERIALS AND METHODS

Bacterial strains and plasmids

Escherichia coli strain DH5 α was used as cloning host strain and *E. coli* strain AD494(DE3) (Novagen) was used as the expression host in this work.

pGEM-T (Promega) was used as cloning vector of PCR products. pJKKm(-) (Kirschman and Cramer, 1988) served as cloning vehicle for fusion protein genes. New, assembled genes for expression were sub-cloned into plasmid pET-11d (Novagen).

Cloning of ANG Δ Cys7Cys236

Restriction endonucleases and DNA modifying enzymes were from New England Biolabs and Promega Corp. Other chemicals and reagents were of analytical reagent grade.

Oligonucleotide primers were synthesised on an Applied Biosystem 394 DNA/RNA Synthesiser using standard phosphoramidite chemistry and were deprotected by heating in ammonium hydroxide solution. Primers were lyophilised and dissolved in TE buffer.

To amplify up a 477 bp long fragment (approximately 2/3 of the whole gene) of C hordein (Accession X60037) from the genomic clone (Lambdaohr1-17) of barley by polymerase chain reaction two specially designed primers were used.

oligonucleotide 1 (SEQ ID NO: 1):

Met Arg Cys
5' GTC ATG AGG CAA CTA AAC CCT TGC AGC CAA GAG TTG CAA TC
BspH I 3'

5 oligonucleotide 2 (SEQ ID NO: 2):

*** Val Cys Gln Gln Pro
5' GGA TCC CTA GAC CAT ACT CCA TAT GCA TGA AGC TTG TTG GGG
BamH I Nde I HindIII
GAC TGG TTG 3'

The reaction was performed in an FTS-4000 Thermal Sequencer (Corbett Research, Australia), with 1 cycle of 3 min at 94°C, 20 s at 55°C, 1 min at 72°C; 36 cycles of 30 s at 94°C, 1 min 30 s at 72°C. The reaction was carried out in 50 µl, containing 45 µl of Supermix (BRL Life Technology) and 1 ng of template DNA and 50 pmol of each oligonucleotides in 5 µl. The DNA was purified following the QIAquick protocol (QIAGEN) and cloned, using the pGEM-T Vector System I (Promega) as was recommended by the manufacturer. White transformant colonies selected for growth in LB medium supplemented with ampicillin (100 mg/l) were screened for insert-bearing plasmid DNA by PCR. Plasmid DNA was purified from positive clones using Jetstar miniprep columns (Genomed) and the insert was sequenced in both direction, using the Prism dye terminator cycle sequencing protocol (Perkin-Elmer). One plasmid isolate, containing the ANG gene, was designated as pGEM-ANG.

This plasmid contained two NdeI restriction sites, one within the gene and one within the pGEM-T multiple cloning site (MCS). To make it unique, the ANG gene was subcloned into pJKKm. Because pJKKm has a HindIII site in MCS, it had to be deleted before subcloning.

One ug of the plasmid was digested with 5 units of HindIII enzyme in 20 µl. Mung Bean Nuclease was used to remove the 5' overhang. Two µl of 10X ZnSO₄ solution and 0.02 units of enzyme was added to the reaction mix and incubated at 30°C. Reaction mix was extracted once with phenol/chloroform after one hour incubation. DNA was recovered with ethanol precipitation and resuspended in 50 µl. An aliquot of 10 µl of the DNA solution was used for ligation. Ligation was carried out in 15 µl solution at 4°C overnight using T4 DNA ligase. The ligation mixture was used to transform competent cells of *E. coli* strain DH5α by electroporation.

Colonies were grown in LB medium supplemented with kanamycin (50 mg/l). Plasmid was purified from three colonies as mentioned above. DNA samples were tested by performing HindIII enzyme digestions. One clone not containing a HindIII site was chosen and designated as pJJK-H.

pJJK-H was cut with SphI and BamHI to subclone pGEM-ANG SphI-BamHI fragment into the plasmid. Both digested DNA were purified on QIAquick columns as described earlier, and ligated in 10 µl solution overnight at 14°C using T4 DNA ligase at a molar ratio of 50:1. Transformed *E. coli* competent cells was spread onto LB medium with kanamycin.

Colonies were tested for insert-bearing plasmid DNA by PCR and 3 positive clones were sequenced as above. One clone was designated as pJANGΔCys7Cys236 and used for further cloning work to assemble genes for fusion proteins. Nucleotide and amino acid sequences of this cloning vehicle are shown in Figure 1.

Design and cloning of oleosin hydrophobic binding domain

The sequence of oleosin hydrophobic binding domain (OHBD) was designed to include the consensus sequence of three (maize, rice and barley) oleosin proteins. The sequence is almost identical to the sequence published for barley oleosin isoform-2 (Aalen 1995, Accession Number X82678). Four primers were designed to encode the protein (Figure 2). The gene was constructed by a modification of the technique of overlap extension, where the two long partially overlapping oligonucleotides were further extended and amplified by shorter external primers, encoding restriction enzyme cleavage sites to clone into pJANGΔCys7Cys236. PCR amplification was performed as above, using the same cycle program and Supermix solution. Concentration of the long oligonucleotides were 0.1 nM, while 2 µM of the short primers in the PCR reaction mix. The PCR product was purified on QIAquick column and cloned into pGEM-T plasmid using Promega kit, as described above. Three positive clones were used for plasmid preparation and sequencing to confirm the nucleotide sequence. One clone was used for further work and called pGEM-OHBD.

Cloning of starch binding domain

The DNA corresponding to the starch binding domain (SBD) of Glucoamylase 1 (1,4-α-D-glucan glucohydrolase) of *Aspergillus niger* (Accession number: X00548) was amplified by PCR from purified genomic

DNA. Primers were designed to allow us to add tails at both ends of the DNA fragment for cloning into pJANGACys7Cys236 vector.

Sequence of the PCR primers are:

oligonucleotide 7 (SEQ ID NO: 3):

5 Gln Ala Cys Thr
5' CAA GCT TGT ACC ACT CCC ACC GCC 3'
Hind III

oligonucleotide 8 (SEQ ID NO: 4):

Ile Cys Arg
10 5' CCA TAT GCA CCG CCA GGT GTC AGT CAC 3'
Nde I

Amplification, cloning into pGEM-T and sequencing was done as described above. One clone bearing the gene fragment was designated as pGEM-SBD.

15 **Cloning of CM16 and CM17 genes for fusion**

Both genes were amplified from purified wheat (*Triticum aestivum*) genomic DNA by PCR.

nucleotide 9 (SEQ ID NO: 5):

5' GTC GGC AAT GAA GAT TGC ACC 3'

20 nucleotide 10 (SEQ ID NO: 6):

5' TCC AAC TGC GTT CTC CTC TTG GCC 3'

nucleotide 11 (SEQ ID NO: 7):

5' GGA TCC CTA GCT CCA CTG AGA CTC 3'

For CM16 gene (accession number X55455) oligonucleotide 10 and 11, while for CM17 gene (accession number X59791) 9 and 11 pairs were used. Clones were called pGEM-CM16 and pGEM-CM17, respectively. For sub-cloning into pJANGACys7Cys236 vector, the genes were PCR amplified again, using pGEM clones as template, purified on QIAquick column and digested with the appropriate enzymes. Primers used in this amplification were:

30 nucleotide 12 (SEQ ID NO: 8):

Gln Ala Leu Gly
5' TGC GCT CAA GCT TTA GGC AAT GAA GAT TGC ACC 3'
Hind III

nucleotide 13 (SEQ ID NO: 9):

Ile Cys Ser

5' CAT ACT CCA TAT GCA GCT CCA CTG AGA CTC 3'

Nde I

5 Cloning puroindoline A gene for fusion

Lambda genomic clone for puroindoline A (PIN-A), kindly provided by Sadequer Rahman, was used as template to amplify the gene (accession number X69913) by PCR. Primers were designed as follows; oligonucleotide 14 (SEQ ID NO: 10):

10 Gln Ala Tyr

5' CAA GCT TAC GAT GTT GCT GGC GGG 3'

Hind III

oligonucleotide 15 (SEQ ID NO: 11):

5' CCA TAT GCA CCA GTA ATA GCC AAT AGT GC 3'

15 Nde I

PCR product was purified, ligated into pGEM-T and sequenced as described above. One clone was used and designated as pGEM-PIN-A.

Fusion of genes or gene fragments with ANG molecule

pJANGΔCys7Cys236 vector was cut with NdeI and HindIII restriction enzymes as all the other pGEM clones and purified on QIAquick columns. Ligation was performed at 14°C overnight in 10 µl of solution containing T4 DNA ligase and insert:vector at about 20:1 molar ratio. One µl of ligation mix was used for transformation of *E. coli* competent cells and spread onto LB plate with kanamycin. Three fusion gene containing colonies were picked up from each transformation for plasmid preparation and sequencing. The clones were called, for example, pJANG/OHBD/Cys7Cys236 in case of the oleosin binding domain containing ANG molecule.

25 The fused genes were subcloned into pET-11d expression vector between NcoI and BamHI sites. These clones were called, for example, pET-ANG/OHBD/Cys7Cys236.

Sequence of PCR primers for detection of genes for ANG/domain/Cys7Cys236 in recombinant wheat

The primers used for the detection of genes for
ANG/domain/Cys7Cys236 in recombinant wheat are set out below.

5

Primer	Length	Sequence	Designation
Bx17_3'	23	CAACCATGTCCTGAACCTTGACC	SEQ ID NO:12
RGS2024	18	TGGCTGTTGAGGTTGCAC	SEQ ID NO:13

Expression of fusion proteins

For protein expression one of the pET clones were transformed into *E. coli* strain AD494(DE3) one day before expression work started.

10

Small scale expression was carried out in 5 ml 2YT medium, supplemented with ampicillin (100 mg/l) using one transformant colony. After about 5 hours from inoculation (OD₆₀₀=0.4), expression of the protein was induced by addition of 0.4 mM isopropyl- β -D-thiogalactopyranoside (IPTG). Both induced and uninduced cultures were further incubated for 4 hours at 37°C. Expression was monitored by SDS-PAGE according to Laemmli (1970).

15

Large scale expression was performed in a shaking flask. One litre 2YT medium was inoculated with 1 ml of overnight culture and induced to express protein by addition of 0.4 mM IPTG at a cell density of ~0.6 Ab. The culture was incubated with shaking overnight and cells were harvested by centrifugation.

20

Detection of proteins

PAGE gels were stained overnight with 0.025% Coomassie Blue R-250 in 10% TCA. Excess stain was washed away by water-ethanol-acetic acid (8:1:1) solution.

25

Immunological detection of the PIN-A containing fusion protein was carried out using the method of Ciaffi et al., (1999). The antibody was raised against puroindoline crude extract, kindly provided by Sadequr Rahman.

Other chimeric proteins were detected in immunoblots using an antibody raised against Hordein, kindly provided by John Skerritt.

30

Purification of the fusion proteins

Expressed ANG/CM16/Cys7Cys236, ANG/CM17/Cys7Cys236 and ANG/SBD/Cys7Cys236 proteins were purified following the method published elsewhere (Tamas et al., 1998), except for the precipitation step. In this work 2 volumes of 1.5 M NaCl were mixed with the 70% ethanol extract, rather than 4 volumes of acetone.

Detection of the gene encoding modified seed storage proteins in transgenic wheat plants

The presence of the gene for modified proteins in transgenic wheat plants was determined by polymerase chain reaction (PCR). Reactions were carried out in 11.6 µl volume containing 9 µl PCR Supermix (GibcoBRL), 50 ng template DNA (extracted from wheat leaf tissue using standard protocols), 172 nmol of each of primers Bx17_3' and RGS2024, and 0.6 µl of 25 mM MgCl₂. The PCR conditions were 1 cycle of 94°C for 2 min; 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 1 min; 1 cycle of 72°C for 4 min, 25°C for 1 min. The results for one plant containing the gene for ANG/OHBD/Cys7Cys236 are shown in Figure 17. The PCR product of approximately 600 base pairs is indicative of the presence of the gene. A 600 bp PCR product is not obtained in PCR reactions from negative control plants.

RESULTS

Design, construction and cloning of pJANGΔCys7Cys236

The gene selected to construct this vector for fusion protein coded for C hordein. This molecule is a storage protein from barley endosperm and characterised by an absence of cysteine residues. The barley genomic clone encodes a molecule of 261 residues, including a 20 residue signal peptide. The gene for mature protein (molecular mass 28 kDa) has 723 nucleotides, including a 669 bp long fragment for a central repetitive domain. The oligonucleotides for PCR were designed to reduce the size of the central part and substitute one residue by cysteine in both unique terminal domains. Oligonucleotide 1 binds to the 5' end of the gene and has an additional sequence at the 5' end to incorporate an initiation ATG codon for methionine and a restriction site for BspHI. Oligonucleotide 1 has a T at position 22 to replace an A to change the codon of serine to cysteine at position 7. The 3' end of oligonucleotide 2 is complimentary to a sequence in the repetitive domain between 430 and 448 nucleotides. This is a rather unique sequence

within the strongly repetitive region and also codes for an end of a repeat motif. The amplified gene contains only a 411 bp (137 amino acids) long fragment for central repetitive domain. This oligonucleotide also contains the whole sequence of the C-terminal unique region (6 amino acids) of C hordein and a restriction site for *Bam*HI, immediately after the stop codon. To change the threonine residue at the position of 236 of the full size molecule (six residues from the C-terminal end) the oligonucleotide has a C at position 26 and an A at position 27 to substitute a G and a T, respectively. There is another base pair change in this primer at position 22 to replace C with A, which is a "wobble" base of the isoleucine codon. This substitution allowed the creation of restriction site for *Nde*I enzyme. Oligonucleotide 2 has also got six nucleotides which are not part of the C hordein gene, coding for alanine and serine. These extra nucleotides were added to create one unique restriction site, close to *Nde*I, within the ANG gene. The two restriction sites are separated with 4 bp to give easy cleavage for both enzymes. The insertion of a gene for another molecule or fragment of a molecule between *Hind*III and *Nde*I enabled the present inventors to create fusion proteins with new, designed characteristics.

The gene for ANG Δ Cys7Cys236 molecule is 474 bp long and codes for a protein with a molecular mass of 18.5 kDa (Figure 3).

PCR amplified DNA was cloned, with two steps, into a plasmid, called pJKKm with a modified MCS. Having deleted *Hind*III restriction site from the original plasmid, the cloning vehicle pJANG Δ Cys7Cys236 has unique *Hind*III and *Nde*I cleavage sites within ANG gene for gene insertion. Size of the vector is 3873 bp and provides kanamycin resistance to host *E. coli* cells. The gene is sitting in the vector with 5' ends close to the SP6 RNA Polymerase transcription initiation site.

Subcloning engineered genes into pET-11d did not require agarose gel purification of the fragments, because of the difference in resistance genes within plasmids.

Expression of ANG Δ Cys7Cys236 protein

To check and compare characteristics of this short molecule to hordeinCys7Cys236, it has been expressed in both small and large quantities. Comparison of the SDS-PAGE patterns of the total cell proteins before and 3 hours after induction with IPTG showed a new band in the induced sample. ANG molecule was readily extracted from lysed cells with 70% (v/v) ethanol

and precipitated by the addition of 2 volume of 1.5 M NaCl solution. The resulting preparation, in the presence of reducing agent (0.1 M DTT), had the same mobility as the extra band in *E. coli* lysate. Apparent molecular mass of the protein was about 21 kDa, having a slightly lower mobility in the gel.

5 This characteristic is not unusual for storage proteins. In the absence of DTT, the ethanol extracted sample gave a ladder of bands indicating that the protein is able to form long chains through disulfide bonds.

The ability of the ANG molecule to incorporate into the gluten matrix of the dough was confirmed by a series of mixing experiments carried out using small scale testers.

10 Results of SDS-PAGE and mixing experiments showed clearly that ANGACys7Cys236 protein had the same or similar properties as the 2 cysteine residue containing C hordein protein.

Design, construction and cloning of synthetic gene for oleosin hydrophobic binding domain (OHBD)

15 Amino acid sequence was designed according to a comparison of four molecules (one maize, two barleys and one rice), using Genetic Computer Group (GCG) program, called "pileup". OHBD gene fragment, for this work, contain the entire region of the lipophilic stretch of oleosin (Figure 2), very similar to barley gene (accession number: X82677). Codon usage was designed to avoid long stretch of Gs and Cs, which could have led to mis-annealing and sequencing problems. The 5' flanking region of the gene fragment consisted of a tripeptide sequence, of the amphipathic N terminus, while the 3' end another tripeptide of the C terminus of oleosin. The central anti-parallel beta stranded domain had 71 residues. The turn of the anti-parallel consists of 13 residues and is the most conserved region. The two anti-parallel strands are highly symmetrical in the pairing of residues of similar hydrophobicity on the opposite strands. These characteristics are very similar to those reported for maize oleosin protein (Huang 1996).

20 The synthetic gene fragment, flanked with two appropriate restriction sites for subcloning, was cloned into pGEM-T vector. One clone, containing the correct sequence of a fragment of 243 bp nucleotides, was designated as pGEM-OHBD. Size of OHBD fragment was 77 amino acids and had a molecular mass of 7 kDa.

Cloning of starch binding domain (SBD)

Glucamylase 1 from *Aspergillus niger* comprises two domains, one being a catalytic domain (1-470 residues) and the other (509-616 residues) being responsible for binding granular starch (Le Gal-Coeffet et al., 1995).

DNA corresponding to the SBD was amplified by PCR method from *Aspergillus niger* purified genomic DNA, cloned into pGEM-T vector, sequenced and designated as pGEM-SBD. This clone had a 337 bp long fragment (Figure 4) with two restriction sites for insertion into pJANGACys7Cys236 vector and an extra alanine before the very first cysteine of SBD. This residue derived from HindIII restriction site. The binding domain had 108 amino acids, including two cysteines, with a molecular mass of 11.9 kDa.

Cloning of CM16 and CM17 gene for fusion

Both CM16 and CM17 (CM refers to Chloroform/Methanol soluble) proteins, reported as members of the α -amylase/trypsin inhibitor family and also reported that specific lipids are tightly bound to the fraction, were purified from wheat (Kobrehel and Sauvaire, 1990). These two molecules are very similar on the amino acid level, but there are a few differences in the distribution of charged residues (Figure 5).

To clone these two genes, three primers were used to amplify them in two, separate PCR reactions from wheat genomic DNA. One primer (oligonucleotide 11) hybridised to the 3' end of the genes, while to distinguish between the two genes, two specific primers were designed for the 5' ends. pGEM-CM17 clone carried the gene for only the mature protein. However, pGEM-CM16 clone had a few extra base pairs from the signal peptide region. Both clones had a DNA fragment encoding mature chloroform/methanol soluble proteins with 10 cysteine residues and a molecular mass of 13.4 kDa.

To subclone these genes into ANGACys7Cys236 carrying vector, two restriction sites were added to one of each ends by PCR, using the same primer pairs for both pGEM clones, as templates. Primer corresponding to the N-terminus of the mature proteins contained nucleotides for HindIII restriction enzyme. It also had an extra alanine residue, and a mutation in the first codon (TTA for leucine), substituting valine (GTC) in CM17 and isoleucine (ATC) in CM16 molecule, because of the sequence requirement for

HindIII enzyme. The fragment cloned into ANG cloning vehicle had 360 nucleotides.

Cloning puroindoline A (PIN-A) gene for fusion

The clone for PIN-A protein, which is capable of binding tightly to both wheat phospholipids and glycolipids (Dubreil et al., 1997), was kindly provided by S. Rahman, derived from a genomic library.

PCR amplified DNA encoded for a 121 amino acid long fragment of puroindoline A, 5 residues shorter in the N terminal region than the mature protein. This shorter protein had 10 cysteine residues and a molecular mass of 14.3 kDa (Figure 6). For insertion into ANG molecule, the pGEM-PIN-A clone also carried two restriction sites (HindIII and NdeI) and an extra alanine, because of HindIII site requirement.

Expression of ANG/OHBD/Cys7Cys236 protein

Having inserted OHBD gene fragment into pJANGΔCys7Cys236 cloning vehicle, the entire gene was subcloned into expression vector pET-11d. Protein was expressed in *E. coli*, using AD494(DE3) strain. This genetically engineered new gene (717 bp long) coded for a protein with a molecular mass of 25.5 kDa.

Cells were harvested by centrifugation after overnight expression (5 ml culture) and resuspended in 100 μ l gel loading buffer (0.125 M TRIS/HCl pH 6.8, 4% SDS, 10% (v/v) glycerol), containing 0.1 M DTT. Proteins were analysed by SDS-PAGE and stained with Coomassie blue (Figure 7). Lane A contains standard molecular weight markers, lane B shows a control in which the host *E. coli* cell contains the pET-11d vector, and lane C shows the expression of the recombinant ANG/OHBD/Cys7Cys236 protein (migrating slightly further than the 30 kDa marker) from pET-11d containing the ANG/OHBD/Cys7Cys236 gene. The apparent molecular mass of the newly synthesised protein (~29 kDa) was larger than that calculated on the basis of the nucleotide sequence of DNA. A similar characteristic has been observed previously for different prolamines of cereals. This fusion protein consisted of part of a prolamine and a part of an oleosin molecule. C hordein had been considered as a rod shaped molecule, while oleosin hydrophobic binding domain has anti-parallel strands, which would penetrate into the oil body and thus anchor the protein stably. It means that none of the two fragments has globular structure as compared with the molecular weight marker molecules. Discrepancy in apparent molecular mass on SDS-PAGE was

probably due to these two structural phenomena and it was manifested in lower mobility. Figure 8 shows a Western blot of an SDS-PAGE gel (run in the absence of reducing agents such as dithiothreitol) using antibodies to the C-hordein protein. Lane A contained standard molecular weight markers, lane B contained extract of an *E. coli* cell expressing a modified C-hordein gene containing a single cysteine in the N-terminal region, and lane C contained extract from *E. coli* containing the ANG/OHBD/Cys7Cys236 gene in the pET-11d expression plasmid. The strong antibody cross reaction to hordein antibodies at 29 kDa confirms that the 29 kDa protein is the product of the ANG/OHBD/Cys7Cys236 gene.

Expression and analysis of ANG/PIN-A/Cys7Cys236 protein

A DNA fragment comprised of ANGΔCys7Cys236, and Puroindoline-A genes with a length of 840 nucleotides from kanamycin resistant plasmid (see above) was subcloned into pET-11d expression vector. One transformant colony was picked up and transferred into 5 ml of 2YT medium. Protein expression was induced by IPTG and after 2 and 6 hours cells were harvested by centrifugation and resuspended in 200 µl of gel loading buffer. SDS-PAGE analysis of protein content of bacteria, before and after induction, is shown on Figure 10. Lane A contained standard protein molecular mass markers. Lane B contained extract of cells harbouring the control pET-11d plasmid, lanes C and D contained extracts of cells harbouring the pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene. Lane C contained extract 2 hours after IPTG induction, lane D contained extract prepared 6 hours following IPTG induction. An extra protein band appeared on gel with an apparent molecular mass of 35 kDa, after induction (lane D). This expressed protein positively reacted with puroindoline A antibody (Figure 11). Lanes A and B contained extract of cells harbouring pET-11d containing the ANG/PIN-A/Cys7Cys236 gene, 2 hours and 6 hours after induction respectively. Lane C contained a Western blot of cells harbouring a plasmid containing the puroindoline A gene alone (not inserted into the ANGΔCys7Cys236 gene), and lane D contained extract of cells harbouring the control pET-11d vector. All lanes were reacted with anti- puroindoline A antibody. Figure 12 shows a Western blot of an SDS-PAGE gel reacted with anti-hordein antibodies. Lane A contained molecular mass markers, lane B contained extract of cells harbouring the control pET-11d vector, lane C contained extract of cells harbouring a plasmid which contains the

puroindoline A gene alone, and lane D contained extract of cells harbouring pET-11d vector containing the ANG/PIN-A/Cys7Cys236 gene. The expressed protein is clearly seen in lane D.

Analogue glutenin-puroindoline A fusion protein had a slightly lower mobility on SDS-PAGE as it was calculated (32.8 kDa) on the basis of amino acid sequence. This small difference was probably due to the rod shape of the ANG part of the molecule. Expressed puroindoline A protein, however, had the same mobility as it was calculated.

Expression of modified proteins following wheat transformation

Microprojectile bombardment is currently the most widely applied technique used to transfer genes into wheat. The transfer of genes into wheat and their expression is carried out using specific DNA constructs containing a selectable marker gene and the gene-of-interest respectively. The selectable marker gene used was contained on the plasmid pEmuKON (Chamberlain et al., 1994). Constructs for the gene-of-interest constructs are shown in Figure 13. Wheat transformation was carried out using paromomycin selection according to the procedures outlined in Witrzens et al., 1998).

Expression, purification and analysis of ANG/SBD/Cys7Cys236 protein

PET11d and the derivative of that plasmid containing the gene for ANG/ SBD/Cys7Cys236 protein (pET-SBD) were separately transformed into strain AD494(DE3). One transformant colony of each was used to inoculate 5 ml cultures of LB media containing 200 mg/l ampicillin. After 4 hours shaking at 37°C these cultures were used to inoculate 600 ml of the same media in 2 l flasks, and these were incubated with shaking overnight (16 hours) at 37°C. IPTG was added to these cultures 6 hours after inoculation to induce expression of protein from the plasmids. Cells from each culture were harvested by centrifugation at 10,000 g for 15 min at 5°C.

One half of the cells from each culture were lysed by incubation for 30 min at room temperature in 35 ml of 50 mM Tris-Cl buffer pH 7.5 containing 8 M urea, 5 mM EDTA and 1 mM DTT. Insoluble material was removed by centrifugation at 2×10^5 g for 1 hr at 15°C. The supernatants were then extensively dialysed against the extraction buffer minus urea (50 mM Tris-Cl pH 7.5, 5 mM EDTA, 1 mM DTT) at 5°C. The cloudy precipitates were removed by centrifugation at 2×10^5 g for 1 hr at 5°C and the supernatants (35 mL) stored at -20°C.

β -Cyclodextrin-binding proteins were purified from 20 ml of each of the two dialysed and clarified bacterial extracts by passage through a 5 ml β -cyclodextrin-sepharose column equilibrated with 50 mM Tris-Cl pH 7.5, 100 mM NaCl, 40 mM dithiothreitol, 1 mM EDTA (column buffer). After loading, the column was washed with 20 ml column buffer and bound proteins eluted with 25 ml column buffer containing 15 mM β -cyclodextrin. Five ml fractions were collected during loading, washing and elution of the column. Fractions were dialysed extensively against water, lyophilised and resuspended in 0.1 M Tris-Cl pH 6.8 at ca. 1 μ l per ml equivalent of the original bacterial cell culture.

Native- and SDS-PAGE gels of the final load fraction and the wash and elution fractions are shown in Figure 14. Each lane contained protein from the equivalent of 7 ml of the bacterial culture. Two major β -cyclodextrin-binding proteins were seen in the elution fractions (see lanes 6 and 7) from the strain bearing plasmid pET-SBD. These are estimated to be 30 kDa and >60 kDa from the SDS-PAGE gel. A 30 kDa protein was also eluted from the column of the extract from the control strain bearing plasmid pET11d (panel C, lane 6), but it is not produced to the same level as in the former strain.

Expression, purification and analysis of ANG/CM16/Cys7Cys236 protein

The gene for this fusion protein (834 bp long) was subcloned, from kanamycin resistant, pJCKm originated plasmid, into the ampicillin resistant pET-11d expression vector, as above. Bacteria was grown at 37°C until OD₆₀₀ reached 0.6 units, then expression was induced by adding IPTG. Cells were harvested two, four or six hours after induction and samples were prepared as in case of SBD chimeric protein.

Result of the SDS-PAGE analysis is shown in Figure 9 (the gel is stained with Coomassie protein stain). Lane A contains standard protein molecular weight markers. Lanes B to E show ethanol-soluble extracts of the crude *E. coli* lysates. Lane B contained extract from cells containing the control plasmid, pET-11d. Lanes C, D and E contained ethanol soluble extracts of cells harbouring the pET-11d vector containing the ANG/CM16/Cys7Cys236 gene, prepared from cells 2, 4 and 6 hours respectively after induction of protein synthesis using IPTG. The ethanol-soluble fraction gave only one band on Coomassie stained gel representing a protein with an apparent molecular mass of 35 kDa. This was also higher

than the calculated (31.9 kDa), and this discrepancy is also explained by the unusual shape of the protein.

Yield of protein expression in large scale, in shaking flask was 30 mg protein from 1 litre of medium.

5 **Demonstration of Starch Binding Activity**

The starch-binding activities of the major proteins in fractions 6 and 7 of the affinity purification column were demonstrated as follows. Fractions were loaded onto four native-PAGE gels containing glycerol (10% w/v), CHAPS (0.05% w/v) and various combinations of β -cyclodextrin (5 mM) and starch (corn amylopectin, 0.1% w/v), as indicated in Figure 18. The mobility of the protein bands in each gel was compared. It was found that the electrophoretic mobility of the major protein band in each fraction was significantly retarded in the presence of starch (Figure 18, panel B). The specificity of the starch binding activity of these protein bands was demonstrated by its competitive inhibition by β -cyclodextrin (panels C, D).

15 **Identification of the purified starch-binding proteins by tryptic-peptide mass fingerprinting and N-terminal protein sequencing**

The two major protein bands from the native gel of β -cyclodextrin affinity column fractions from strain AD494(DE3)/pET-SBD were subjected to in-gel tryptic digestion followed by matrix-assisted laser-desorption time-of-flight (MALDI-TOF) mass spectroscopy. The masses of the identified tryptic fragments were compared to the theoretical masses of tryptic peptides expected from ANG/SBD/Cys7Cys236, as well as being used to search the protein sequence databases for other candidate proteins giving matching tryptic peptides.

The 30 kDa protein of fraction 6 (Figure 14, panels A, B, lane 6) was identified as the periplasmic maltose-binding protein of *E. coli* (product of the *malE* gene), based on 17 peptide matches covering 57% of the protein sequence.

The major protein band of fraction 7 (Figure 14, panels A, B, lane 7) was identified as ANG/SBD/Cys7Cys236 on the basis of 4 matching peptides covering 23.11% of the sequence (Table 1). Further proof of the identity of this protein was obtained from N-terminal protein sequencing. Ten cycles of sequencing yielded the sequence:

35 M R Q L N P C S Q E

which is as expected for ANG/SBD/Cys7Cys236. This provides convincing proof of the identity of the purified protein, and confirms that the functional data are properties of the modified seed storage protein, as designed.

5 Table 1. Mass spectral characterisation of ANG/SBD/Cys7Cys236

Peptide mass	Expected mass	Start	End	Sequence
1776.74	1777.28	224	238	IESDDSVWESEDPNR
2193.00	2193.85	221	238	FIRIESDDSVWESEDPNR
2652.26	2653.54	198	220	YTSSDPLWYVTVTLPAGESFEYK
2257.03	2258.00	239	258	EYTVPQAC*GTSTATVTDTWR

C* = acrylamide adduct of cysteine

10 **Demonstration of the ability of ANG/SBD/Cys7Cys236 to form disulphide-bonded polymers**

The design of the analogue glutenin protein (ANG) included cysteine residues at both N- and C-termini to allow disulphide-mediated incorporation of the modified seed-storage protein into the gluten macropolymer. The present inventors have shown by *in vitro* polymerisation experiments that the ability of the ANG/domain/Cys7Cys236 proteins to incorporate into disulphide-bonded polymeric species is not impaired by the presence of the extra domain.

20 The ability to form polymers was demonstrated by copper-mediated oxidation of ANG/SBD/Cys7Cys236 protein in 50 mM Tris-Cl buffer pH 6.8. Three pairs of protein samples were employed, with one of each pair containing added β -cyclodextrin (0.5 mM) to analyse any effect of a ligand on the disulphide-bonding of the ANG. Two pairs were oxidised for 5 min at room temperature in the presence of 0.1 mM CuSO₄, before chelation of the metal by addition of EDTA (10 mM). One of these pairs was then reduced in the presence of 20 mM DTT. A further pair of samples had no copper added, but were reduced with DTT in the presence of EDTA. The samples were analysed by SDS-PAGE (Figure 19). There was no effect on oxidation due to the presence of the ligand β -cyclodextrin. The majority of the oxidised protein sample forms a high molecular weight band which just enters the separating gel (lanes 1, 2). That this change in mobility is due to the

formation of disulphide-bond polymeric forms is shown by the ability of the disulphide-specific reductant DTT (lanes 3, 4) to break the polymer back down into the higher-mobility form seen in the reduced sample (lanes 5, 6).

Detection of the gene encoding modified seed storage proteins in transgenic wheat plants.

The presence of the gene for modified proteins in transgenic wheat plants was determined by polymerase chain reaction (PCR). Reactions were carried out in 11.6 μ l volume containing 9 μ l PCR Supermix (GibcoBRL), 50 ng template DNA (extracted from wheat leaf tissue using standard protocols), 172 nmol of each of primers Bx17_3' and RGS2024, and 0.6 μ l of 25 mM $MgCl_2$. The PCR conditions were 1 cycle of 94°C for 2 min; 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 72°C for 1 min; 1 cycle of 72°C for 4 min, 25°C for 1 min. The results for one plant containing the gene for ANG/OHBD/Cys7Cys236 are shown in Figure 17. The PCR product of approximately 600 base pairs is indicative of the presence of the gene. A 600 bp PCR product is not obtained in PCR reactions from negative control plants.

Demonstration of the ability of ANG/SBD/Cys7Cys236 to form disulphide-bonded polymers

The design of the analogue glutenin protein (ANG) included cysteine residues at both N- and C-termini to allow disulphide-mediated incorporation of the modified seed-storage protein into the gluten macropolymer. The present inventors have shown by *in vitro* polymerisation experiments that the ability of the ANG/domain/Cys7Cys236 proteins to incorporate into disulphide-bonded polymeric species is not impaired by the presence of the extra domain.

The ability to form polymers was demonstrated by copper-mediated oxidation of the protein in 50 mM Tris-Cl buffer pH 6.8. Three pairs of protein samples were employed, with one of each pair containing added β -cyclodextrin (0.5 mM) to analyse any effect of ligand-binding on the disulphide-bonding of the ANG moiety. Two pairs were oxidised for 5 min at room temperature in the presence of 0.1 mM $CuSO_4$, before chelation of the metal by addition of EDTA (10 mM). One of these pairs was then reduced in the presence of 20 mM DTT. A further pair of samples had no copper added, but were reduced with DTT in the presence of EDTA. The samples were analysed by SDS-PAGE (Figure19). There was no effect on oxidation due to

the presence of the ligand β -cyclodextrin. The majority of the oxidised protein sample forms a high molecular weight band which just enters the separating gel (lanes 1, 2). That this change in mobility is due to the formation of disulphide-bond polymeric forms is shown by the ability of the disulphide-specific reductant DTT (lanes 3, 4) to break the polymer back down into the higher-mobility form seen in the reduced samples (lanes 5, 6).
SUMMARY

The demonstration of the binding activity of a modified glutenin according to the present invention was provided in the example given above for what has been termed the ANG/SBD/Cys7Cys236 protein. The data provided supports the demonstration of the expression and purification of a novel protein (apparent molecular mass 60 kDa) from the *E. coli* strain bearing the expression plasmid (pET-SBD), in comparison to the lack of this protein in the same strain bearing the control plasmid (pET11d) (Figure 14). The purification involved the use of an affinity chromatography step in which the protein was purified on the basis of its ability to bind to a column matrix containing β -cyclodextrin. The protein was specifically eluted from this column using unbound β -cyclodextrin. This, therefore, represents the first example of how the present inventors demonstrate reduction to practice through the generation of a seed storage protein containing a functional ligand binding domain.

The demonstration of the binding of the major purified proteins to corn starch using native polyacrylamide gel electrophoresis, and the inhibition of this binding by β -cyclodextrin (PAGE) (Figure18) is a highly convincing demonstration of reduction to practice of the present invention.

The identification of this protein as ANG/SBD/Cys7Cys236 by peptide mass fingerprinting and N-terminal protein sequencing further demonstrates the functionality of modified seed-storage proteins containing a macromolecular binding domain. The anomolous apparent molecular weight of this protein in SDS-PAGE may be yet another example of the general phenomenon of modified C hordein proteins migrating with unexpected electrophoretic mobility, as the calculated mass of the protein is only 30 kDa. Alternatively, the protein may form a dimer, perhaps around a single bound β -cyclodextrin ligand, which is not dissociated by boiling in the presence of SDS and a reducing agent.

The ANG portion forms disulphide-bonded polymeric species was demonstrated by native-PAGE of the copper-oxidised affinity purified proteins. The fusion protein has been demonstrated to contain the novel starch binding activity. Importantly, it is believed that the modified protein has not lost the ability to polymerise through the cysteine residues at the N- and C-termini and therefore the novel protein can be incorporated into the gluten macropolymer in wheat flour.

Demonstration of the capacity of the modified seed-storage protein with macromolecule-binding domain (in this instance ANG/SBD/Cys7Cys236) to form disulphide-bonded polymeric species (Figure 19). This demonstrates the ability of the protein to be incorporated into the gluten macropolymer through disulphide bonds.

Table 1 demonstrates the identity between the molecular weights of the experimentally determined tryptic fragments of the major β -cyclodextrin-binding protein (starch-binding protein) purified from the ANG/SBD/Cys7Cys236 strain and the theoretical masses of the tryptic fragments of the ANG/SBD/Cys7Cys236 protein. The identification of the 30kDa protein as an endogenous bacterial maltose-binding protein is also provided.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

REFERENCES

- Aalen, R.B. 1995. The transcript encoding two oleosin isoforms are both present in the aleurone and in the embryo of barley (*Hordeum vulgare*) seeds. *Plant Mol. Biol.* 28:583-588
- Bekes, F., Gras, P.W. and Gupta, R.B. 1994. Mixing properties as a measure of reversible reduction and oxidation of doughs. *Cereal Chemistry* 71:44-50
- Bushuk, W. 1998. Interactions in wheat doughs. *In* Interactions: The keys to cereal quality. R.J. Hamer, R.C. Hoseney Eds. American Association of Cereal Chemists, Inc. St Paul, Minnesota. pp1-16.
- Chamberlain, D.A., Brettell, R.I.S., Last, D.I., Witrzens, B., McElroy, D., Dolferus, R., Dennis, E.S. 1994. The use of the EMu promoter with antibiotic and herbicide resistance genes for the selection of transgenic wheat callus and rice plants. *Australian Journal of Plant Physiology*. 21: 95-112.
- Ciaffi, M., Lee, Y.K., Tamas, L., Gupta, R., Skerritt, J. and Appels, R. 1999. The low-molecular-weight glutenin subunit proteins of primitive wheats. III. The genes from D-genome species. *Theoretical and Applied Genetics* 98: 135-148.
- Dubreil, L., Compoin, J-P. and Marion, D. 1997. Interaction of Puroindolines with wheat flour polar lipids determines their foaming properties. *J. Agric. Food Chem.* 45:108-116
- Gan, Z., Ellis, P.R. and Schofield, J.D. 1995. Gas cell Stabilisation and gas retention in wheat bread dough. *Journal of Cereal Science* 21:215-230
- Huang, A.H.C. 1996. Oleosins and oil bodies in seeds and other organs. *Plant Physiology* 110:1055-1061
- Kasarda, D.D. 1989. Glutenin structure in relation to wheat quality. Pages 277-302 in : *Wheat is unique*. Y. Pomeranz, Ed. Amer. assoc. Cereal Chem., St Paul, MN.
- Kirschman, J.A. and Cramer, J.H. 1988. Two new tools: multi-purpose cloning vectors that carry kanamycin or spectinomycin/streptomycin resistance markers. *Gene* 68: 163-165.
- Kobrehel, K. and Sauvaire, Y. 1990. Particular lipid composition in isolated proteins of durum wheat. *J. Agric. Food Chem.* 38:1164-1171
- Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685

Le Gal-Coeffet, M-F., Jacks, A.J., Sorimachi, K., Williamson, M.P., Williamson, G. and Archer, D.B. 1995. Expression in *Aspergillus Niger* of the starch-binding domain of glucoamylase. *Eur. J. Biochem.* 233:561-567

MacRitchie, F. 1992. Physicochemical properties of wheat proteins in relation to functionality. *Adv. Food Nutr. Research* 36:1-87

Morrison, W.R. 1989. Recent progress on the chemistry and functionality of flour lipids. Pages 131-149 in: *Wheat end-use properties: Wheat and flour characterisation for specific end-uses*. H. Salovaara, Ed. University of Helsinki, Lahti.

Tamas, L., Bekes, F., Greenfield, J., Tatham, A.S., Gras, P.W., Shewry, P.R. and Appels, R. 1998. Heterologous expression and dough mixing studies of wild-type and mutant C hordeins. *Journal of Cereals Science* 27:15-22

Weeks, J.T., Anderson, O.D. and Blechl, A.E. 1993. Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*). *Plant Physiology* 102:1077-1084

Witzens, B., Brettell, R.I.S., Murray, F.R., McElroy, D., Li, Z., Dennis, E.S. 1998. Comparison of three selectable marker genes for transformation of wheat by microprojectile bombardment. *Australian Journal of Plant Physiology* 25:39-44.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Commonwealth Scientific and Industrial Research Organisation

(B) STREET: Limestone Avenue

(C) CITY: Campbell

(D) STATE: ACT

(E) COUNTRY: Australia

(F) POSTAL CODE (ZIP): 2601

(A) NAME: Goodman Fielder Limited

(B) STREET: Level 42 Grosvenor Place

(C) CITY: Sydney

(D) STATE: NSW

(E) COUNTRY: Australia

(F) POSTAL CODE (ZIP): 2000

(A) NAME: Groupe Limagrain Pacific Pty Ltd

(B) STREET: Level 31, 1 O'Connell Street

(C) CITY: Sydney

(D) STATE: NSW

(E) COUNTRY: Australia

(F) POSTAL CODE (ZIP): 2000

(ii) TITLE OF INVENTION: Modified Proteins

(iii) NUMBER OF SEQUENCES: 26

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GTCATGAGGC AACTAAACCC TTGCAGCCAA GAGTTGCAAT C

41

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 51 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGATCCCTAG ACCATACTCC ATATGCATGA AGCTTGTGG GGGACTGGTT G

51

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CAAGCTTGTA CCACTCCAC CGCC

24

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CCATATGCAC CGCCAGGTGT CAGTCAC

27

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GTCGGCAATG AAGATTGCAC C

21

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TCCAACTGCG TTCTCCTCTT GGCC

24

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGATCCCTAG CTCCTAGAG ACTC

24

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TGCCTCAAG CTTTAGGCAA TGAAGATTGC ACC

33

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CATACTCCAT ATGCAGCTCC ACTGAGACTC

30

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CAAGCTTACG ATGTTGCTGG CGGG

24

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCATATGCAC CAGTAATAGC CAATAGTGC

29

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CAACCATGTC CTGAACCTTC ACC

23

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TGGCTGTTGA GGTGACAC

18

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ATAGAATACA GCATGCTCCC GGCGGCCATG GCCGCGGGAT TGTCATGAGG CAACTAAACC 60
CTTGCAGCGT CCCCCAACAA GCTTCATGCA TATGGAGTAT GGTCTAGGGA TCCGGGTACC 120
GAGCTCGAAT TCGCCCTATA 140

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Met Arg Gln Leu Asn Pro Cys Ser Val Pro Gln Gln Ala Ser Cys Ile
1 5 10 15
Trp Ser Met Val
20

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 243 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

```

CAAGCTTTGA AAGCCGCTAC TGCACAGCA GCTGGATCAT TGCTTGTGCT ATCCGGACTA 60
ATACTAGCTG GCACAGTCAT AGCACTACA GTGGCCACAC CAGTGCTAGT CATATTTAGC 120
CCAGTGCTAG TGCCAGCGGC CATAGCCCTA GCGCTAATGT CAGCAGGCTT TGTCACGTCA 180
GGCGGGCTGG GCGTGGGTGC GCTGAGCTCC TTTAGTGTGT TAGCCAATAC TGCTGCATA 240
TGG

```

243

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```

Gln Ala Leu Lys Ala Ala Thr Ala Thr Ala Ala Gly Ser Leu Leu Val
 1             5             10             15
Leu Ser Gly Leu Ile Leu Ala Gly Thr Val Ile Ala Leu Thr Val Ala
          20             25             30
Thr Pro Val Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Ala Ile
      35             40             45
Ala Leu Ala Leu Met Ser Ala Gly Phe Val Thr Ser Gly Gly Leu Gly
      50             55             60
Val Ala Ala Leu Ser Ser Phe Ser Val Leu Ala Asn Thr Ala Cys Ile
      65             70             75             80
Trp

```

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 477 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

```

ATGAGGCAAC TAAACCTTG CAGCCAAGAG TTGCAATCAC CACAACAATC ATATCTGCCG 60
CAGCCATATC CACAAAACCC ATATCTACCG CAAAACCAT TTCCAGTGCA GCAACCGTTT 120
CACACACCCC AACAAATATT CCCCTATCTA CCAGAGGAAT TGTTCCTCCA ATATCAAATA 180
CCAACCCCCC TACAACCACA ACAACCATT CCCCACAAC CACAACAACC TCTTCCTCGG 240
CCCCAACAAC CATTCCCCTG GCAACCACAA CAACCATTTT CCCAGCCCCA AGAACCAATT 300
CCCCAACAAC CATTCCCCTG GCAACCACAA CAACCATTTT CCCAGCCCCA AGAACCAATT 360
CAACAATAAA TTTTCCAGCA ACCCCAACAA TCATACCCTG TGCAACCTCA ACAGGCCATT 420
CCTCAACAAC CTCAACCACT CCCCCAACA GCTTCATGCA TATGGAGTAT GGTCTAG 477

```

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 158 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

```

Met Arg Gln Leu Asn Pro Cys Ser Gln Glu Leu Gln Ser Pro Gln Gln
 1           5           10          15
Ser Tyr Leu Gln Gln Pro Tyr Pro Gln Asn Pro Tyr Leu Pro Gln Lys
          20          25          30
Pro Phe Pro Val Gln Gln Pro Phe His Thr Pro Gln Gln Tyr Phe Pro
          35          40          45
Tyr Leu Pro Glu Glu Leu Phe Pro Gln Tyr Gln Ile Pro Thr Pro Leu
          50          55          60
Gln Pro Gln Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Leu Pro Arg
          65          70          75          80
Pro Gln Gln Pro Phe Pro Trp Gln Pro Gln Gln Pro Phe Pro Gln Pro
          85          90          95

```

Gln Glu Pro Ile Pro Gln Gln Pro Gln Gln Pro Phe Pro Gln Gln Pro
 100 105 110

Gln Gln Pro Phe Pro Gln Gln Pro Gln Gln Ile Ile Phe Gln Gln Pro
 115 120 125

Gln Gln Ser Tyr Pro Val Gln Pro Gln Gln Pro Phe Pro Gln Gln Pro
 130 135 140

Gln Pro Val Pro Gln Gln Ala Ser Cys Ile Trp Ser Met Val
 145 150 155

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 338 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AAGCTTCTAC CACTCCCACC GCCGTGGCTG TGACTTTCGA TCTGACAGCT ACCACCACCT 60
 ACGGCGAGAA CATCTACCTG GTCGGATCGA TCTCTCAGCT GGGTGACTGG GAAACCAGCG 120
 ACGGCATAGC TCTGAGTGCT GACAAGTACA CTTCCAGCGA CCCGCTCTGG TATGTCACTG 180
 TGACTCTGCC GGCTGGTGAG TCGTTTGAGT ACAAGTTTAT CCGCATTGAG AGCGATGACT 240
 CCGTGGAGTG GGAGAGTGAT CCCAACCGAG AATACACCGT TCCTCAGGCG TGCGGAACGT 300
 CGACCGCGAC GGTGACTGAC ACCTGGCGGT GCATATGG 338

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 112 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Ala Ser Thr Thr Pro Thr Ala Val Ala Val Thr Phe Asp Leu Thr Ala
 1 5 10 15

Thr Thr Thr Tyr Gly Glu Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln
 20 25 30

Leu	Gly	Asp	Trp	Glu	Thr	Ser	Asp	Gly	Ile	Ala	Leu	Ser	Ala	Asp	Lys
	35						40					45			
Tyr	Thr	Ser	Ser	Asp	Pro	Leu	Trp	Tyr	Val	Thr	Val	Thr	Leu	Pro	Ala
	50					55					60				
Gly	Glu	Ser	Phe	Glu	Tyr	Lys	Phe	Ile	Arg	Ile	Glu	Ser	Asp	Asp	Ser
	65				70					75					80
Val	Glu	Trp	Glu	Ser	Asp	Pro	Asn	Arg	Glu	Tyr	Thr	Val	Pro	Gln	Ala
				85					90				95		
Cys	Gly	Thr	Ser	Thr	Ala	Thr	Val	Thr	Asp	Thr	Trp	Arg	Cys	Ile	Trp
			100					105					110		

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 371 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

AAGCTTTCGG	CAATGAAGAT	TGCACCCCAT	GGATGAGTAC	TCTGATCACT	CCACTCCCAA	60
GCTGCCGTGA	CTATGTGGAA	CAACAAGCAT	GTGCGATCGA	AACGCCCGGG	TCGCCGTACC	120
TCGCCAAGCA	GCAGTGCTGT	GGGGAGCTTG	CAAACATTCC	GCAGCAGTGC	CGATGCCAGG	180
CGTGCGCTA	CTTCATGGGG	CCGAAGTCTC	GTCCGGATCA	GAGCGGCCTC	ATGGAAGTCC	240
CCGGATGCCC	TAGGGAGGTG	CAGATGGACT	TCGTGAGGAT	ACTCGTCACG	CCGGGGTACT	300
GCAACTTGAC	GACCGTTCAC	AACACTCCGT	ACTGCCTCGC	TATGGAGGAG	TCTCAGTGGA	360
GCTGCATATG	G					371

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

```

Ala Ile Gly Asn Glu Asp Cys Thr Pro Trp Met Ser Thr Leu Ile Thr
 1           5           10
Pro Leu Pro Ser Cys Arg Asp Tyr Val Glu Gln Gln Ala Cys Arg Ile
          20           25           30
Glu Thr Pro Gly Ser Pro Tyr Leu Ala Lys Gln Gln Cys Cys Gly Glu
          35           40           45
Leu Ala Asn Ile Pro Gln Gln Cys Arg Cys Gln Ala Leu Arg Tyr Phe
          50           55           60
Met Gly Pro Lys Ser Arg Pro Asp Gln Ser Gly Leu Met Glu Leu Pro
65           70           75           80
Gly Cys Pro Arg Glu Val Gln Met Asp Phe Val Arg Ile Leu Val Thr
          85           90           95
Pro Gly Tyr Cys Asn Leu Thr Thr Val His Asn Thr Pro Tyr Cys Leu
          100          105          110
Ala Met Glu Glu Ser Gln Trp Ser Cys Ile Trp
          115          120

```

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 123 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

```

Ala Ile Gly Asn Glu Asp Cys Thr Pro Trp Thr Ser Thr Leu Ile Thr
 1           5           10
Pro Leu Pro Ser Cys Arg Asn Tyr Val Glu Glu Gln Ala Cys Arg Ile
          20           25           30
Glu Met Pro Gly Pro Pro Tyr Leu Ala Lys Gln Glu Cys Cys Glu Gln
          35           40           45
Leu Ala Asn Ile Pro Gln Gln Cys Arg Cys Gln Ala Leu Arg Tyr Phe
          50           55           60
Met Gly Pro Lys Ser Arg Pro Asp Gln Ser Gly Leu Met Glu Leu Pro
65           70           75           80
Gly Cys Pro Arg Glu Val Gln Met Asn Phe Val Pro Ile Leu Val Thr
          85           90           95

```

Pro Gly Tyr Cys Asn Leu Thr Thr Val His Asn Thr Pro Tyr Cys Leu
 100 105 110

Gly Met Glu Glu Ser Gln Trp Ser Cys Ile Trp
 115 120

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 377 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

AAGCTTACGA TGTTGCTGGC GGGGGTGGTG CTCACAATG CCCTGTAGAG ACAAAGCTAA 60
 ATTCATGCAG GAATTACCTG CTAGATCGAT GCTCAACGAT GAAGGATTTC CCGGTCACCT 120
 GGCGTTGGTG GAAATGGTGG AAGGGAGGTT GTCAAGAGCT CCTTGGGGGAG TGTGTCAGTC 180
 GGCTCGGCCA AATGCCACCG CAATGCCGCT GCAACATCAT CCAGGGGTCA ATCCAAGGCG 240
 ATCTCGGTGG CATCTTCGGA TTTCAGCGTG ATCGGGCAAG CAAAGTGATA CAAGAAGCCA 300
 AGAACCTGCC GCCCAGGTGC AACCAGGGCC CTCCTGCAA CATCCCCGGC ACTATTGGCT 360
 ATTACTGGTG CATATGG 377

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 125 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Ala Tyr Asp Val Ala Gly Gly Gly Gly Ala Gln Gln Cys Pro Val Glu
 1 5 10 15
 Thr Lys Leu Asn Ser Cys Arg Asn Tyr Leu Leu Asp Arg Cys Ser Thr
 20 25 30
 Met Lys Asp Phe Pro Val Thr Trp Arg Trp Trp Lys Trp Trp Lys Gly
 35 40 45

Gly	Cys	Gln	Glu	Leu	Leu	Gly	Glu	Cys	Cys	Ser	Arg	Leu	Gly	Gln	Met
50						55					60				
Pro	Pro	Gln	Cys	Arg	Cys	Asn	Ile	Ile	Gln	Gly	Ser	Ile	Gln	Gly	Asp
65					70					75					80
Leu	Gly	Gly	Ile	Phe	Gly	Phe	Gln	Arg	Asp	Arg	Ala	Ser	Lys	Val	Ile
				85					90					95	
Gln	Glu	Ala	Lys	Asn	Leu	Pro	Pro	Arg	Cys	Asn	Gln	Gly	Pro	Pro	Cys
			100					105					110		
Asn	Ile	Pro	Gly	Thr	Ile	Gly	Tyr	Tyr	Trp	Cys	Ile	Trp			
		115					120					125			

100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000
 1001
 1002
 1003
 1004
 1005
 1006
 1007
 1008
 1009
 1010
 1011
 1012
 1013
 1014
 1015
 1016
 1017
 1018
 1019
 1020
 1021
 1022
 1023
 1024
 1025
 1026
 1027
 1028
 1029
 1030
 1031
 1032
 1033
 1034
 1035
 1036
 1037
 1038
 1039
 1040
 1041
 1042
 1043
 1044
 1045
 1046
 1047
 1048
 1049
 1050
 1051
 1052
 1053
 1054
 1055
 1056
 1057
 1058
 1059
 1060
 1061
 1062
 1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072
 1073
 1074
 1075
 1076
 1077
 1078
 1079
 1080
 1081
 1082
 1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133
 1134
 1135
 1136
 1137
 1138
 1139
 1140
 1141
 1142
 1143
 1144
 1145
 1146
 1147
 1148
 1149
 1150
 1151
 1152
 1153
 1154
 1155
 1156
 1157
 1158
 1159
 1160
 1161
 1162
 1163
 1164
 1165
 1166
 1167
 1168
 1169
 1170
 1171
 1172
 1173
 1174
 1175
 1176
 1177
 1178
 1179
 1180
 1181
 1182
 1183
 1184
 1185
 1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449

CLAIMS:

1. A method for producing a modified glutenin or seed-storage protein, the method comprising adding to the protein an exogenous amino acid domain which confers to the modified protein the ability to bind a ligand or other macromolecule, wherein the modified protein has an ability to incorporate into gluten.
2. The method according to claim 1 wherein the domain includes one or more cysteine residues.
3. The method according to claim 1 wherein one or more cysteine residues are incorporated at one or both ends of the amino acid sequence of the protein.
4. The method according to claim 1 wherein the modified protein further comprises one or more exogenous amino acid residues which confer to the protein an enhanced ability to incorporate into gluten.
5. The method according to claim 1 wherein the ligand or other macromolecule to which the domain binds is selected from the group consisting of lipid, and starch.
6. The method according to claim 5 wherein the domain capable of binding lipid is derived from barley oleosin gene or the lipid-binding regions of wheat CM16 protein.
7. The method according to claim 5 wherein the domain capable of binding starch is derived from glucoamylase from *Aspergillus niger*.
8. The method according to any one of claims 1 to 7 wherein the glutenin or seed-storage protein is selected from the group consisting of low molecular weight glutenins, high molecular weight glutenins, gliadins, puroindolines, grain softness proteins, friabilins, and Chloroform/Methanol-soluble proteins.
9. The method according to claim 8 wherein the glutenin or seed-storage protein is C hordein from barley.
10. A modified glutenin or seed-storage protein having an ability to incorporate into gluten and having an exogenous amino acid domain inserted therein which confers to the modified protein the ability to bind a ligand or other macromolecule produced by the method according to any one of claims 1 to 9.

11. A modified glutenin or seed-storage protein having an ability to incorporate into gluten and having an exogenous amino acid domain inserted therein which confers to the modified protein the ability to bind a ligand or other macromolecule.
12. A modified glutenin or seed-storage protein according to claim 11 wherein the ligand or other macromolecule to which the domain binds is selected from the group consisting of lipid, and starch.
13. A modified glutenin or seed-storage protein according to claim 12 wherein the domain capable of binding lipid is derived from barley oleosin gene or the lipid-binding regions of wheat CM16 protein.
14. A modified glutenin or seed-storage protein according to claim 12 wherein the domain capable of binding starch is derived from glucoamylase from *Aspergillus niger*.
15. A modified glutenin or seed-storage protein according to any one of claims 10 to 14 wherein the glutenin or seed-storage protein is selected from the group consisting of low molecular weight glutenins, high molecular weight glutenins, gliadins, puroindolines, grain softness proteins, friabilins, and Chloroform/Methanol-soluble proteins.
16. A modified glutenin or seed-storage protein selected from the group consisting of ANG/SBD/Cys7Cys236, ANG/OHBD/Cys7Cys236, and ANG/CM16/Cys7Cys236, as hereinbefore defined.
17. An isolated nucleic acid molecule encoding a modified glutenin or seed-storage protein according to any one of claims 10 to 16.
18. A cell containing an isolated nucleic acid molecule according to claim 17 such that on expression of the nucleic acid molecule, the cell produces the modified glutenin or seed-storage protein.
19. The cell according to claim 18 selected from the group consisting of bacteria, yeast, plant, insect, and mammal.
20. The cell according to claim 19 being *Escherichia coli*.
21. The cell according to claim 19 being *Pichia* sp. or *Saccharomyces cerevisiae*.
22. The cell according to claim 19 being a recombinant wheat cell.
23. Use of a modified glutenin or seed-storage protein according to any one of claims 10 to 16 in the preparation of a food product.

24. The use according to claim 23 wherein the food product is selected from the group consisting of leavened or unleavened breads, pasta, noodles, breakfast cereals, snack foods, cakes, pastries, and foods containing flour-based sauces.

5 25. Use of a modified glutenin or seed-storage protein according to any one of claims 10 to 16 in the preparation of a non-food product.

26. The use according to claim 25 wherein the non-food product is selected from the group consisting of films, coatings, adhesives, building materials, and packaging materials.

10 27. Use of a grain or part of a grain containing a modified glutenin or seed-storage protein according to any one of claims 10 to 16 in the preparation of a food product.

PCT/AU99/00563
Received 25 July 2000

-16

ABSTRACT

A method for producing a modified glutenin or seed-storage protein, the method comprising adding to the protein an exogenous amino acid domain which confers to the modified protein the ability to bind a ligand or other macromolecule, wherein the modified protein has an ability to incorporate into gluten, modified glutenin or seed-storage proteins, and uses thereof.

09745533 100004

AMERICAN CHEMICAL
SOCIETY

5'... ATAGATAC A GCATGC TCC CGCGCG CCATGG CGCGCG GATTGTC ATG AGG CAA CTA NAC CCT TGC AGC....
 SF6 Promoter<| SphI EagI NcoI SacII BspHI M R Q L N P C S
 V P Q Q A s C I W S M V ***
 ANG SEQUENCE....GTC CCC CAA CAA ect TCA TGC ATA TGG AGT ATG GTC TAG GGATCC
 HindIII NdeI BamHI

GGTACC GAGCTC GAATTC GCCTATA... 3'
 KpnI SacI EcoRI |> T7 Promoter

Figure 1

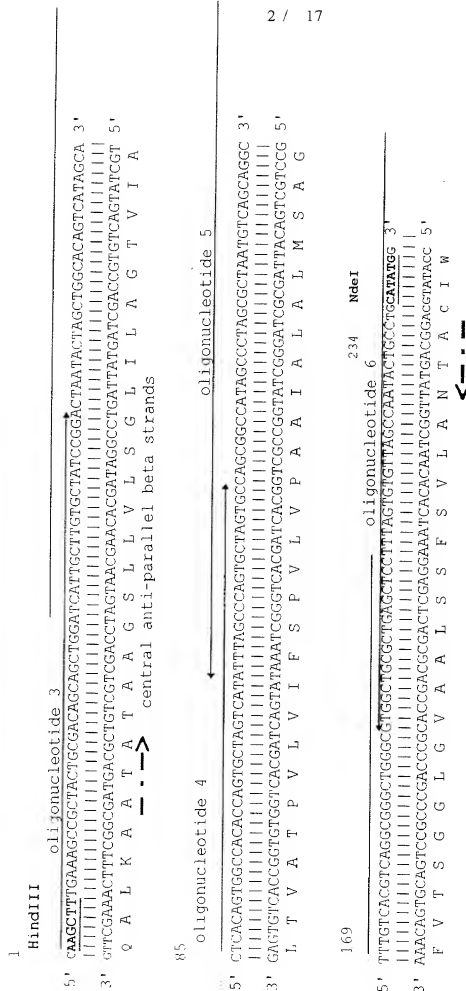


Figure 2

3 / 17

1
ATGAGGCAACTAAACCTT**TGC**AGCCAAGAGTTGCAATCACCACAACAATCATATCTGCCG
M R Q L N P **Cys** S Q E L Q S P Q Q S Y L Q

61
CAGCCATATCCACAAAACCCATATCTACCGCAAAAACCATTTCCAGTGCAGCAACCGTTT
Q P Y P Q N P Y L P Q K P F P V Q Q P F

121
CACACACCCCAACAATATTTCCCTATCTACCAGAGGAATTGTTTCCCAATATCAAATA
H T P Q Q Y F P Y L P E E L F P Q Y Q I

181
CCAACCCCCCTACAACCACAACAACCATTTCCCCAACAACCACAACAACCTCTTCCTCGG
P T P L Q P Q Q P F P Q Q P Q Q P L P R

241
CCCCAACAACCATTTCCCTGGCAACCACAACAACCATTTCCCAGCCCCAAGAACCAATT
P Q Q P F P W Q P Q Q P F P Q P Q E P I

301
CCCCAGCAACCACAACAACCATTTCCCACAGCAACCACAACAACCATTTCCCACAGCAACCA
P Q Q P Q Q P F P Q Q P Q Q P F P Q Q P

361
CAACAAATAATTTTCCAGCAACCCCAACAATCATACCTGTGCAACCTCAACAGCCATTT
Q Q I I F Q Q P Q Q S Y P V Q P Q Q P F

421 477
CCTCAACACCTCAACCAAGTCCCCCAACAA GCT TCA **TGC**ATATGGAGTATGGTCTAG
P Q Q P Q P V P Q Q A S **Cys** I W S M V ***

Figure 3

4 / 17

HindIII 1 54
AAGCTTCTACCACTCCCACCGCCGTGGCTGTGACTTTTCGATCTGACAGCTACCACCACCTAC
A S T T P T A V A V T F D L T A T T T Y

114
GGCGAGAACATCTACCTGGTCGGATCGATCTCTCAGCTGGGTGACTGGGAAACCAGCGAC
G E N I Y L V G S I S Q L G D W E T S D

174
GGCATAGCTCTGAGTGCTGACAAGTACACTTCCAGCGACCCGCTCTGGTATGTCACTGTG
G I A L S A D K Y T S S D P L W Y V T V

234
ACTCTGCCGGCTGGTGAGTCGTTTGAGTACAAGTTTATCCGCATTGAGAGCGATGACTCC
T L P A G E S F E Y K F I R I E S D D S

294
GTGGAGTGGGAGAGTGATCCCAACCGAGAATACACCGTTCTCAGGCGTGCGGAACGTCG
V E W E S D P N R E Y T V P Q A C G T S

321 NdeI
ACCGCGACGGTGACTGACACCTGGCGGTGCATATG
T A T V T D T W R C I W

Figure 4

5 / 17

HindIII

57

AAGCTTTCGGCAATGAAGATTGCACCCCATGGATGAGTACTCTGATCACTCCACTCCCAAGCA I G N E D C T P W M S T L I T P L P S
CM17. T

117

TGCCGTGACTATGTGGAACAACAAGCATGTGCGATCGAAACGCCGGGTGCGCGTACCTC

C R D Y V E Q Q A C R I E T P G S P Y L
. . N . . . E M . . P . . .

177

GCCAAGCAGCAGTGCTGTGGGGAGCTTGCAAACATTCGCGAGCAGTGCCGATGCCAGGCG

A K Q Q C C G E L A N I P Q Q C R C Q A
. . . E . . E Q

237

CTGCGCTACTTCATGGGGCCGAAGTCTCGTCCGGATCAGAGCGGCTCATGGAACCCCC

L R Y F M G P K S R P D Q S G L M E L P
.

297

GGATGCCCTAGGGAGGTGCAGATGGACTTCGTGAGGATACTCGTCACGCCGGGGTACTGC

G C P R E V Q M D F V R I L V T P G Y C
. N . . P

354

AAGTTGACGACCGTTACAACTCCGTACTGCCTCGCTATGGAGGAGTCTCAGTGG

N L T T V H N T P Y C L A M E E S Q W
. G

357 NdeI

AGCTGCATATGG

S C I W

Figure 5

6 / 17

HindIII 57
AAGCTTACGATGTTGCTGGCGGGGGTGGTGTCAACAATGCCCTGTAGAGACAAAGCTAAAT
A Y D V A G G G G A Q Q C P V E T K L N

117
TCATGCAGGAATTACCTGCTAGATCGATGCTCAACGATGAAGGATTTCCCGGTACCTGG
S C R N Y L L D R C S T M K D F P V T W

177
CGTTGGTGGAAATGGTGAAGGGAGGTTGTCAAGAGCTCCTTGGGGAGTGTTCAGTCGG
R W W K W W K G G C Q E L L G E C C S R

237
CTCGGCCAAATGCCACCGCAATGCCGCTGCAACATCATCCAGGGGTCAATCCAAGGCGAT
L G Q M P P Q C R C N I I Q G S I Q G D

297
CTCGGTGGCATCTTCGGATTTTCAGCGTGATCGGGCAAGCAAAGTGATACAAGAAGCCAAAG
L G G I F G F Q R D R A S K V I Q E A K

300
AACCTGCCGCCAGGTGCAACCAGGGCCCTCCCTGCAACATCCCCGGCACTATTGGCTAT
N L P P R C N Q G P P C N I P G T I G Y

363 NdeI
TACTGGTGCCATATGG
Y W C I W

Figure 6

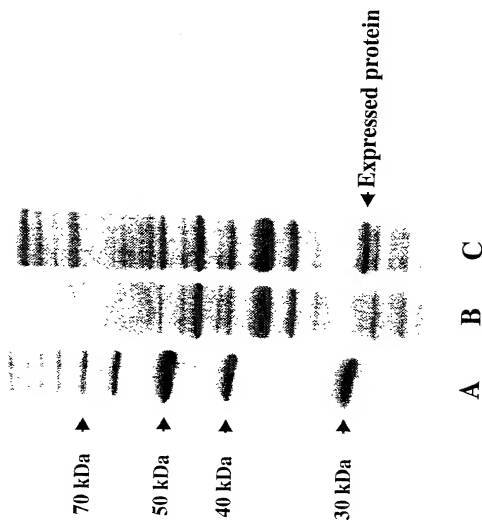


Figure 7

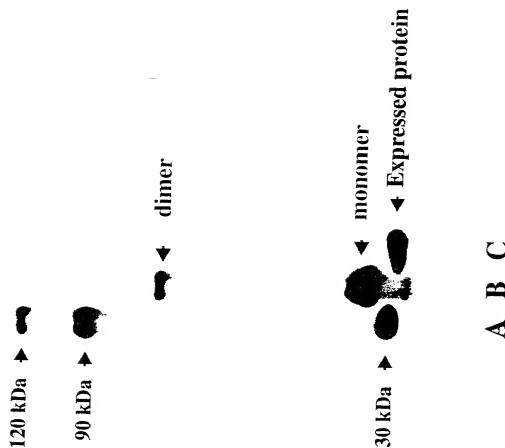


Figure 8

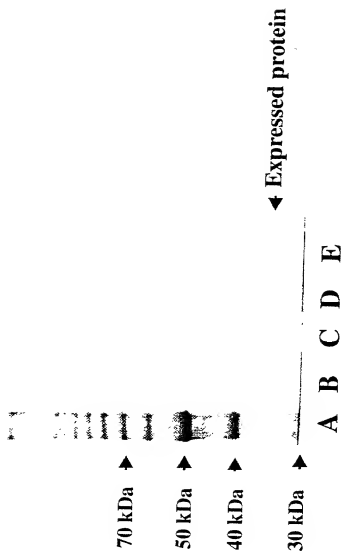


Figure 9

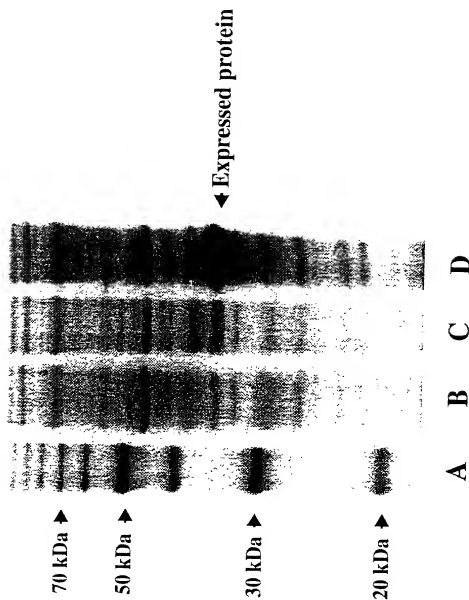


Figure 10

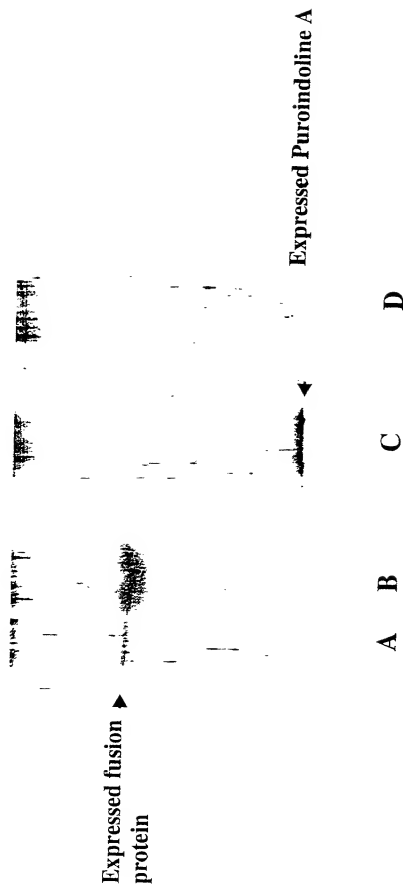


Figure 11

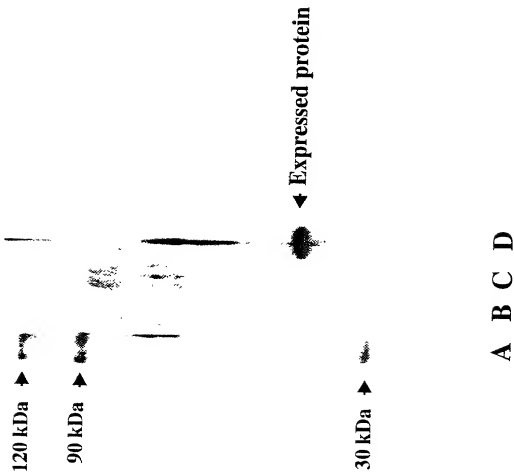


Figure 12.

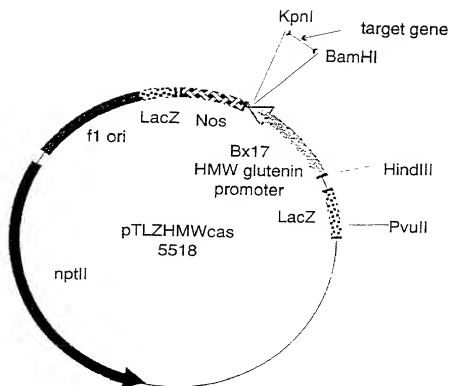


Figure 13

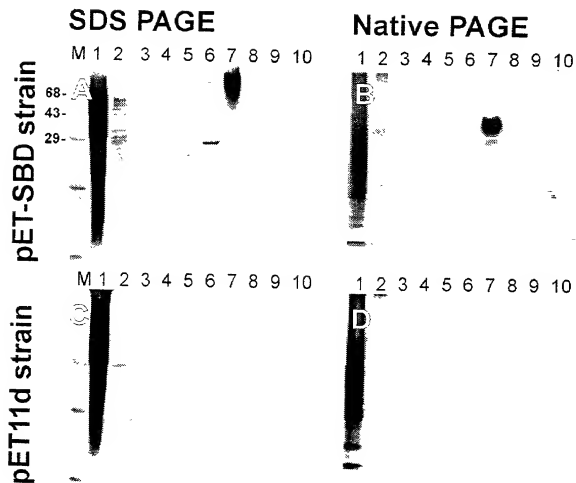


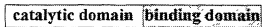
Figure 14

15 / 17

ANGCys7Cys236



binding domain in another protein



ANG/insert/Cys7Cys236

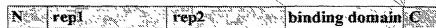


Figure 15

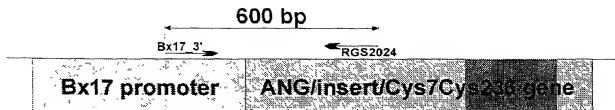


Figure 16

16 / 17

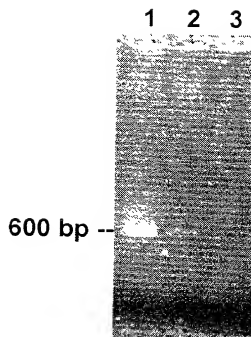


Figure 17

Fraction: 6 7 6 7 6 7 6 7
A B C D



Starch: - + - +
 β -CD: - - + +

Figure 16

17 / 17

	M	1	2	3	4	5	6
oxidised		+	+	+	+	-	-
reduced		-	-	+	+	+	+
β -CD	markers	+	-	+	-	+	-



Figure 19

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: MODIFIED PROTEINS the specification of which

☐

is attached hereto

☒

was filed on:

July 12, 1999

as Application Serial No.:

PCT/AU99/000563

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the Patent Office all information known to me to be material to patentability as defined in 37 C.F.R. 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s):

Priority Claimed

PP 4604
(Number)

Australia
(Country)

10 July 1998
(Day/Month/Year Filed)

☒ ☐

Yes No

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose to the Patent Office all information known to me to be material to patentability as defined in 37 C.F.R. 1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

(Application Serial No.)

(Filing Date)

(Status: patented, pending, abandoned)

I hereby appoint the following attorneys to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Harold C. Hohbach, Reg. No. 17,757; Aldo J. Test, Reg. No. 18,048; Donald N. MacIntosh, Reg. No. 20,316; Edward S. Wright, Reg. No. 24,903; David J. Brezner, Reg. No. 24,774; Robert B. Chickering, Reg. No. 24,286; Richard F. Trecartin, Reg. No. 31,801; Steven F. Caserza, Reg. No. 29,780; Michael A. Kaufman, Reg. No. 32,988; Edward N. Bachand, Reg. No. 37,085; R. Michael Ananian, Reg. No. 35,050; Robin M. Silva, Reg. No. 38,304; Maria S. Swiatek, Reg. No. 37,244; Todd A. Lorenz, Reg. No. 39,754, provided that if any one of said attorneys ceases being affiliated with the law firm of Flehr Hohbach Test Albritton & Herbert LLP as partner, employee or of counsel, such attorney's appointment as attorney and all powers derived therefrom shall terminate on the date such attorney ceases being so affiliated.

Direct all telephone calls to Richard F. Trecartin at (415) 494-8700.

Address all correspondence to:

FLEHR HOHBACH TEST
ALBRITTON & HERBERT LLP
Suite 3400, Four Embarcadero Center
San Francisco, California 94111

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Title 18, United States Code, §1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-00
Full name of sole or first inventor:

Rudi APPELS

Inventor's signature:

Date:

11-07-01

Residence:

Australia

Citizenship:

Australia

Post Office Address:

40 Gingana Street, Aranda, ACT 2614, Australia

14X

2-00
Full name of second inventor:

Matthew MORELL

Inventor's signature:

Date:

11-07-01

Residence:

Australia

Citizenship:

Australia

Post Office Address:

33 Wangara Street, Aranda, ACT 2614, Australia

14X

3-00
Full name of third inventor:

Frank BEKES

Inventor's signature:

Date:

13-8-01

Residence:

Australia

Citizenship:

Australia

Post Office Address:

34 Hull Road, Beecroft, NSW 2119, Australia

14X

4-00
Full name of fourth inventor:

Laszlo TAMAS

Inventor's signature:

Date:

4 Oct 2001

Residence:

Hungary

Citizenship:

Hungary

Post Office Address:

Valvog utca, 9, H-1033 Budapest, Hungary

14X

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Commonwealth Scientific and Industrial
Research Organisation

(B) STREET: Limestone Avenue

(C) CITY: Campbell

(D) STATE: ACT

(E) COUNTRY: Australia

(F) POSTAL CODE (ZIP): 2601

(A) NAME: Goodman Fielder Limited

(B) STREET: Level 42 Grosvenor Place

(C) CITY: Sydney

(D) STATE: NSW

(E) COUNTRY: Australia

(F) POSTAL CODE (ZIP): 2000

(A) NAME: Groupe Limagrain Pacific Pty Ltd

(B) STREET: Level 31, 1 O'Connell Street

(C) CITY: Sydney

(D) STATE: NSW

(E) COUNTRY: Australia

(F) POSTAL CODE (ZIP): 2000

(ii) TITLE OF INVENTION: Modified Proteins

(iii) NUMBER OF SEQUENCES: 26

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GTGATGAGGC AACTAAACCC TTGCAGCCAA GAGTGTCAAT C

41

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGATCCCTAG ACCATACTCC ATATGCATGA AGCTTGTGG GGGACTGGTT G

51

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CAAGCTTGTA CCACTCCAC CGCC

24

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CCATATGCAC CGCCAGGTGT CAGTCAC

27

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GTCGGCAATG AAGATTGCAC C

21

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TCCAAC TCGC TTCTCCTCTT GGCC

24

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGATCCCTAG CTCCACTGAG ACTC

24

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TGCGCTCAAG CTTTAGGCAA TGAAGATTGC ACC

33

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CATACTCCAT ATGCAGCTCC ACTGAGACTC

30

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CAAGCTTACG ATGTTGCTGG CGGG

24

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCATATGCAC CAGTAATAGC CAATAGTGC

29

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CAACCATGTC CTGAACCTTC ACC

23

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TGGCTGTTGA GGTTGCAC

18

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```
ATAGAATACA GCATGCTCCC GGGCGCCATG GCCGCGGGAT TGTCATGAGG CAACTAAACC 60
CTTGACGCGT CCCCCAACAA GCTTCATGCA TATGGAGTAT GGTCTAGGGA TCCGGGTACC 120
GAGCTCGAAT TCGCCCTATA                                     140
```

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

```
Met Arg Gln Leu Asn Pro Cys Ser Val Pro Gln Gln Ala Ser Cys Ile
 1           5           10           15
Trp Ser Met Val
                20
```

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 243 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

```
CAAGCTTTGA AAGCCGCTAC TGCGACAGCA GCTGGATCAT TGCTTGTGCT ATCCGGACTA 60
ATACTAGCTG GCACAGTCAT AGCACTCACA GTGGCCACAC CAGTGTAGT CATATTTAGC 120
CCAGTGCTAG TGCCAGCGGC CATAGCCCTA GCGCTAATGT CAGCAGGCTT TGTCACGTCA 180
```

GGCGGGCTGG GCGTGGCTGC GCTGAGCTCC TTTAGTGTGT TAGCCAATAC TGCCTGCATA 240
TGG 243

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Gln	Ala	Leu	Lys	Ala	Ala	Thr	Ala	Thr	Ala	Ala	Gly	Ser	Leu	Leu	Val
1			5						10					15	
Leu	Ser	Gly	Leu	Ile	Leu	Ala	Gly	Thr	Val	Ile	Ala	Leu	Thr	Val	Ala
			20					25					30		
Thr	Pro	Val	Leu	Val	Ile	Phe	Ser	Pro	Val	Leu	Val	Pro	Ala	Ala	Ile
			35				40					45			
Ala	Leu	Ala	Leu	Met	Ser	Ala	Gly	Phe	Val	Thr	Ser	Gly	Gly	Leu	Gly
			50			55					60				
Val	Ala	Ala	Leu	Ser	Ser	Phe	Ser	Val	Leu	Ala	Asn	Thr	Ala	Cys	Ile
			65			70				75				80	

Trp

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 477 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

ATGAGGCAAC TAAACCCTTG CAGCCAAGAG TTGCAATCAC CACAACAATC ATATCTGCCG 60
CAGCCATATC CACAAAACCC ATATCTACCG CAAAAACCAT TTCCAGTGCA GCAACCGTTT 120
CACACACCCC AACAATATTT CCCCTATCTA CCAGAGGAAT TGTTCCTCCA ATATCAATA 180
CCAACCCCCC TACAACCACA ACAACCATTC CCCCACAAC CACAACAACC TCTTCCTCGG 240
CCCCACAAC CATTCCTCTG GCAACCACAA CAACCATTC CCCAGCCCCA AGAACCAATT 300

CCCCAACAAC CATTCCCCTG GCAACCACAA CAACCATTTT CCCAGCCCCA AGAACCAATT 360
 CAACAAATAA TTTTCCAGCA ACCCCAACAA TCATACCCTG TGCAACCTCA ACAGCCATT 420
 CCTCAACAAC CTCACCAGT CCCCCAACAA GCTTCATGCA TATGGAGTAT GGTCTAG 477

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 158 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Met	Arg	Gln	Leu	Asn	Pro	Cys	Ser	Gln	Glu	Leu	Gln	Ser	Pro	Gln	Gln
1				5				10						15	
Ser	Tyr	Leu	Gln	Gln	Pro	Tyr	Pro	Gln	Asn	Pro	Tyr	Leu	Pro	Gln	Lys
			20					25					30		
Pro	Phe	Pro	Val	Gln	Gln	Pro	Phe	His	Thr	Pro	Gln	Gln	Tyr	Phe	Pro
			35				40					45			
Tyr	Leu	Pro	Glu	Glu	Leu	Phe	Pro	Gln	Tyr	Gln	Ile	Pro	Thr	Pro	Leu
			50			55					60				
Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	Pro	Arg
					70					75					80
Pro	Gln	Gln	Pro	Phe	Pro	Trp	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro
					85					90				95	
Gln	Glu	Pro	Ile	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro
			100					105					110		
Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Ile	Ile	Phe	Gln	Gln	Pro
			115				120						125		
Gln	Gln	Ser	Tyr	Pro	Val	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro
			130				135					140			
Gln	Pro	Val	Pro	Gln	Gln	Ala	Ser	Cys	Ile	Trp	Ser	Met	Val		
			145			150				155					

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 338 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

```

AAGCTTCTAC CACTCCCACC GCGGTGGCTG TGACTTTTGA TCTGACAGCT ACCACCACCT  60
ACGGCGAGAA CATCTACCTG GTCGGATCGA TCTCTCAGCT GGGTGACTGG GAAACCAGCG 120
ACGGCATAGC TCTGAGTGCT GACAAGTACA CTTCCAGCGA CCCGCTCTGG TATGTCAGTG 180
TGACTCTGCC GGCTGGTGAG TCGTTTGAGT ACAAGTTTAT CCGCATTGAG AGCGATGACT 240
CCGTGGAGTG GGAGAGTGAT CCCAACCAGG AATACACCGT TCCTCAGGCG TCGGGAACGT 300
CGACCGCGAC GGTGACTGAC ACCTGGCGGT GCATATGG                               338

```

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 112 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

```

Ala Ser Thr Thr Pro Thr Ala Val Ala Val Thr Phe Asp Leu Thr Ala
1           5           10           15
Thr Thr Thr Tyr Gly Glu Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln
20          25          30
Leu Gly Asp Trp Glu Thr Ser Asp Gly Ile Ala Leu Ser Ala Asp Lys
35          40          45
Tyr Thr Ser Ser Asp Pro Leu Trp Tyr Val Thr Val Thr Leu Pro Ala
50          55          60
Gly Glu Ser Phe Glu Tyr Lys Phe Ile Arg Ile Glu Ser Asp Asp Ser
65          70          75          80
Val Glu Trp Glu Ser Asp Pro Asn Arg Glu Tyr Thr Val Pro Gln Ala
85          90          95
Cys Gly Thr Ser Thr Ala Thr Val Thr Asp Thr Trp Arg Cys Ile Trp
100         105         110

```

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 371 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

```

AAGCTTTCGG CAATGAAGAT TGCACCCCAT GGATGAGTAC TCTGATCACT CCACTCCCAA 60
GCTGCCGTGA CTATGTGGAA CAACAAGCAT GTCGCATCGA AACGCCGGG TCGCCGTACC 120
TCGCCAAGCA GCAGTGTCTGT GGGGAGCTTG CAAACATTCC GCAGCAGTGC CGATGCCAGG 180
CGCTGCGCTA CTTTCATGGGG CCGAAGTCTC GTCCGGATCA GAGCGGCCTC ATGGAAGTCC 240
CCGGATGCCC TAGGGAGGTG CAGATGGACT TCGTGAGGAT ACTCGTACAG CCGGGGTACT 300
GCAACTTGAC GACCGTTCAC AACACTCCGT ACTGCCTCGC TATGGAGGAG TCTCAGTGGA 360
GCTGCATATG G 371

```

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 123 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

```

Ala Ile Gly Asn Glu Asp Cys Thr Pro Trp Met Ser Thr Leu Ile Thr
1           5           10           15
Pro Leu Pro Ser Cys Arg Asp Tyr Val Glu Gln Gln Ala Cys Arg Ile
20           25           30
Glu Thr Pro Gly Ser Pro Tyr Leu Ala Lys Gln Gln Cys Cys Gly Glu
35           40           45
Leu Ala Asn Ile Pro Gln Gln Cys Arg Cys Gln Ala Leu Arg Tyr Phe
50           55           60
Met Gly Pro Lys Ser Arg Pro Asp Gln Ser Gly Leu Met Glu Leu Pro
65           70           75           80
Gly Cys Pro Arg Glu Val Gln Met Asp Phe Val Arg Ile Leu Val Thr
85           90           95
Pro Gly Tyr Cys Asn Leu Thr Thr Val His Asn Thr Pro Tyr Cys Leu
100          105          110
Ala Met Glu Glu Ser Gln Trp Ser Cys Ile Trp
115          120

```

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 123 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

```

Ala Ile Gly Asn Glu Asp Cys Thr Pro Trp Thr Ser Thr Leu Ile Thr
 1           5           10           15
Pro Leu Pro Ser Cys Arg Asn Tyr Val Glu Glu Gln Ala Cys Arg Ile
          20           25           30
Glu Met Pro Gly Pro Pro Tyr Leu Ala Lys Gln Glu Cys Cys Glu Gln
          35           40           45
Leu Ala Asn Ile Pro Gln Gln Cys Arg Cys Gln Ala Leu Arg Tyr Phe
          50           55           60
Met Gly Pro Lys Ser Arg Pro Asp Gln Ser Gly Leu Met Glu Leu Pro
 65           70           75           80
Gly Cys Pro Arg Glu Val Gln Met Asn Phe Val Pro Ile Leu Val Thr
          85           90           95
Pro Gly Tyr Cys Asn Leu Thr Thr Val His Asn Thr Pro Tyr Cys Leu
          100          105          110
Gly Met Glu Glu Ser Gln Trp Ser Cys Ile Trp
          115          120

```

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 377 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

```

AAGCTTACGA TGTGCTGGC GGGGGTGGTG CTCAACAATG CCCTGTAGAG ACAAAGCTAA 60
ATTCATGCAG GAATTACCTG CTAGATCGAT GCTCAACGAT GAAGGATTTC CCGGTCACCT 120
GGCGTTGGTG GAAATGGTGG AAGGGAGGTT GTCAGAGCT CCTTGGGGAG TGTTCAGTC 180
GGCTCGGCCA AATGCCACCG CAATGCCGCT GCAACATCAT CCAGGGGTCA ATCCAAGGCG 240
ATCTCGGTGG CATCTTCGGA TTTCAGCGTG ATCGGGCAGG CAAAGTGATA CAAGAAGCCA 300

```

AGAACCTGCC GCCCAGGTGC AACCAGGGCC CTCCTTGCAA CATCCCCGGC ACTATTGGCT 360
 ATTACTGGTG CATATGG 377

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 125 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Ala	Tyr	Asp	Val	Ala	Gly	Gly	Gly	Gly	Ala	Gln	Gln	Cys	Pro	Val	Glu	1	5	10	15
Thr	Lys	Leu	Asn	Ser	Cys	Arg	Asn	Tyr	Leu	Leu	Asp	Arg	Cys	Ser	Thr	20	25	30	
Met	Lys	Asp	Phe	Pro	Val	Thr	Trp	Arg	Trp	Trp	Lys	Trp	Trp	Lys	Gly	35	40	45	
Gly	Cys	Gln	Glu	Leu	Leu	Gly	Glu	Cys	Cys	Ser	Arg	Leu	Gly	Gln	Met	50	55	60	
Pro	Pro	Gln	Cys	Arg	Cys	Asn	Ile	Ile	Gln	Gly	Ser	Ile	Gln	Gly	Asp	65	70	75	80
Leu	Gly	Gly	Ile	Phe	Gly	Phe	Gln	Arg	Asp	Arg	Ala	Ser	Lys	Val	Ile	85	90	95	
Gln	Glu	Ala	Lys	Asn	Leu	Pro	Pro	Arg	Cys	Asn	Gln	Gly	Pro	Pro	Cys	100	105	110	
Asn	Ile	Pro	Gly	Thr	Ile	Gly	Tyr	Tyr	Trp	Cys	Ile	Trp				115	120	125	